Characterization of Thg1-Like Proteins in tRNA Biogenesis

A Senior Honors Thesis

Presented in partial fulfillment of the requirements for graduation *with research distinction* in Biochemistry in the undergraduate colleges of The Ohio State University

By

Fuad Mohammad

The Ohio State Universiry
May 2012

Project Advisor: Dr. Jane E. Jackman, Department of Chemistry and Biochemistry

Abstract

Transfer RNAs (tRNAs) function to translate the nucleic acid code of messenger RNA into amino acid sequences, and the highly conserved secondary and tertiary structures of tRNAs are critical for proper recognition by the translational machinery. In all domains of life, the evolutionary pressure to maintain tRNA structure has led to diverse pathways in tRNA biogenesis that make use of a 3'-to-5' RNA polymerization reaction catalyzed by the $tRNA^{His}$ guanylyltransferase (Thg1) enzyme family. Here, we study two processes catalyzed by Thg1 enzymes. First, members of the Thg1 enzyme family, designated as Thg1-like proteins (TLPs), have been suggested to participate in a multistep tRNA editing pathway in several lower eukarya. Organisms such as *Dictyostelium discoideum* and *Acanthamoeba castellanii* encode in their mitochondria aberrant tRNAs with mismatches at the normally base-paired tRNA acceptor stem. Two TLPs in *D. discoideum* (DdiTLP3-4) have robust *in vitro* 3'-to-5' polymerization on seven truncated tRNA intermediates proposed to be edited in *D. discoideum*, and the sequence of one tRNA product reaffirms the previously observed preference for templated nucleotide addition. Although DdiTLP3 and DdiTLP4 share similar tRNA repair activities, their mechanisms for repairing truncated tRNAs differ. DdiTLP3 is more sensitive to tRNA length and can only incorporate nucleotides to the first three positions of the 5' end of tRNAs. In contrast, DdiTLP4 repairs larger 5'-end truncations than DdiTLP3, yet it also catalyzes the incorporation of excess nucleotides at the 5' end under sub- physiological ATP concentrations. The second process we studied is the catalysis of G_{-1} addition to $tRNA^{His}$, a process thought to occur in all Eukarya and many Archaea. In *A. castellanii*, however, the two encoded TLPs (AcaTLP1/2) have very weak *in vitro* G_{-1} addition activity, and are likely involved in tRNA editing only. In studying *D. discoideum* and *A. castellanii* Thg1 enzymes, the activities and

mechanistic properties of 3'-5' RNA polymerization provide insight in the evolution of the Thg1 superfamily and the roles they have adopted for tRNA biogenesis.

Introduction

In the study of biology, certain themes often arise from observations that are considered to be universalities. Of these themes, the flow of genetic information from DNA to messenger RNA (mRNA) to a final polypeptide product has withstood evolutionary time to become a ubiquitous feature of life, the central dogma of molecular biology. Processes within this transfer of information and the macromolecules involved are therefore highly conserved amongst all species. Translation, the process by which information stored mRNA is relayed into a functional polypeptide, requires precise interplay between dozens of macromolecules that have co-evolved to recognize each other and the mRNA message. Transfer RNAs (tRNA) exemplify the intricacies of substrate recognition during the translational process, since they have to be recognized non-specifically by the ribosome, and very selectively by their cognate synthetase. To bind to the ribosome, all tRNA species have evolved a common L-shaped tertiary conformation. However, each tRNA must be individually recognized by aminoacyl-tRNA synthetases and charged with the correct amino acid. To solve this problem, synthetases and tRNAs have co-evolved to establish a series of identity elements to distinguish correct aminoacyl-tRNAs from incorrect pairings while maintaining almost identical tRNA tertiary structures.

Like the synthetase, enzymes such as RNase P, Ef-Tu, and many tRNA modifying enzymes have evolved distinct mechanisms for recognizing and acting on tRNAs. This theme of co-evolution can be extended another player in tRNA maturation, the $tRNA^{His}$ guanylyltransferase (Thg1) superfamily. The first member of the Thg1 superfamily was discovered in *Saccharomyces cerevisiae* (yeast Thg1 or yThg1), where it catalyzes the non-templated addition of a guanylate residue at the -1 position (G_{-1}) of yeast cytoplasmic $tRNA^{His}$ (Gu et al., 2003). Unlike most tRNAs, which begin at the +1 position after RNase P cleavage, mature $tRNA^{His}$ contains an

additional G_{-1} residue either base paired across from C_{73} (Bacteria and Archaea), or unpaired across A_{73} (Eukarya) (Figure 2). Excluding a small group of alpha-proteobacteria lacking a G_{-1} (Wang et al., 2007; Yuan et al., 2011), the G_{-1} residue is a distinguishing feature of $tRNA^{His}$ and functions as the primary identity element for HisRS in all systems investigated. In most Eukarya and some Archaea and Bacteria, Thg1 proteins play an essential role in the $tRNA^{His}$/HisRS relationship by catalyzing addition of the G_{-1} identity element (Figure 1A). However, Thg1-catalyzed G_{-1} addition is not ubiquitous, and many bacterial and archaeal species lack a Thg1 protein capable of catalyzing the eukaryotic G_{-1} addition activity (Jackman et al., 2012). It has been observed in these organisms that G_{-1} is genomically encoded and possibly retained after a non-canonical RNase P 5'-end cleavage pattern (Figure 2A) (Holm and Krupp, 1992; Marck and Grosjean, 2002). Interestingly, the phylogenetic relationship between organisms that genomically encode and retain G_{-1} (mostly Archaea and Bacteria) and those that obtain G_{-1} by Thg1 enzymes (most Eukarya and only some Archaea/Bacteria) is reflected in the Thg1 superfamily, which can be subdivided by sequence into a eukaryal branch and archaeal/bacterial branch (Figure 1A). Eukaryal Thg1 enzymes, or true Thg1 enzymes, catalyze the essential G_{-1} addition during maturation of the nuclear-encoded $tRNA^{His}$; while archaeal/bacterial members referred to as Thg1-like proteins (TLPs) catalyze a more general tRNA repair reaction (Abad et al., 2011; Rao et al., 2011). The phylogenetic distribution of Thg1 family members suggest that the eukaryal Thg1 enzymes represent specializations of their archaeal and bacterial counterparts, and have evolved to maintain the G_{-1} requirement of $tRNA^{His}$ in eukaryotes.

Several significant findings provide hints as to the possible ancestry of Thg1 enzymes. First is the unusual 3'-to-5' polymerizing capability of the Thg1 superfamily. In *S. cerevisiae*, yeast Thg1 is highly specific for addition to the cytoplasmic $tRNA^{HisGUG}$, which normally contains an A_{73}

discriminator nucleotide. However, when provided with a C_{73}-containing $tRNA^{His}$ *in vitro*, yeast Thg1 catalyzes 3'-to-5' templated polymerization across C_{74} and C_{75} (Jackman and Phizicky, 2006b). Furthermore, the primary *in vitro* activity of TLPs has been observed to be 3'-to-5' polymerization on truncated 5' tRNA ends in a repair-like reaction, very similar to the polymerization reaction catalyzed by yThg1 (Abad et al., 2011; Rao et al., 2011). The physiological role for this truncated tRNA repair activity is not fully understood, and it has been proposed that TLPs may function as quality control enzymes or compete with tRNA degradation during cellular stress (Rao et al., 2011). It has also been hypothesized that a subset of TLPs catalyzing 5'-tRNA repair may function in mitochondrial tRNA editing, which will be the focus of this study (Abad et al., 2011).

Furthermore, the 3'-to-5' polymerization reaction catalyzed by yThg1 and TLPs is the first of its kind, and characterization of the Thg1 superfamily may provide necessary evidence to understand the evolutionary relationship between 3'-to-5' and 5'-to-3' polymerization. Through time, the 5'-to-3' directionality has become overwhelmingly favored in biology, although the benefits of unidirectional nucleotide synthesis are unclear. For instance, a 3'-to-5' polymerization activity would have eliminated the need for semi-discontinuous lagging strand synthesis during DNA replication, and would altogether circumvent the requirement for telomerase activity since DNA could be replicated to the ends of the chromosome. Synthesis of nucleotide polymers in both directionalities shares almost identical mechanisms. Thg1 enzymes catalyze nucleotidyl transfer by first activating the 5'-monophosphorylated tRNA end using ATP (generating ApptRNA), and subsequently catalyzing nucleophilic attack on this adenylylated intermediate using the 3' hydroxyl of an incoming nucleotide. yThg1 also catalyzes the removal of pyrophosphate from the 5'-end of the growing polynucleotide chain, an activity thought to

prevent additional nucleotide transfers since the 5'-triphosphate resulting from 3'-to-5' addition of each NTP serves as the activated end for subsequent additions during 3'-to-5' polymerization (Figure 2C). Canonical 5'-to-3' polymerases share similar chemistry, particularly at the nucleotidyl transfer step that is at the heart of Thg1 3'-to-5' addition activity**.** Some have argued the energetic cost of proofreading favors nucleic acid synthesis in the 5'-to-3' direction, since proofreading in the 3'-to-5' direction would require reactivation with ATP to restart synthesis of the chain after removal of an incorrect nucleotide. Nevertheless, the energetic costs of lagging strand synthesis in the 5'-to-3' direction outweigh those benefits due to the requirement for ATP in each ligation reaction needed to join Okazaki fragments during replication. Surprisingly, the crystal structure of Thg1 from *Homo sapiens* (human Thg1 or hThg1) reveals structural similarities with the Pol A family of DNA and RNA polymerases, including a predicted two-metal ion active site (Hyde et al., 2010). The structural relationship between Thg1 enzymes and canonical DNA and RNA polymerases suggests the possibility of shared ancestry between conventional 5'-to-3' polymerization and the 3'-to-5' activities of the Thg1 superfamily.

Here, we focus on advancing the current understanding of the evolution of the Thg1 family and how these proteins have influenced tRNA biogenesis. We also discuss how Thg1 members have adapted to their roles, including a detailed comparison of the mechanisms employed by both Thg1 enzymes and TLPs that have allowed them to use 3'-to-5' polymerization for tRNA-specific reactions, and how the presence of Thg1 enzymes have led to alternative methods of tRNA processing, such as tRNA editing and repair. To address these issues, we focused on the Thg1 homologues of the slime mold *Dictyostelium discoideum*, and of the amoeboid protozoan A*canthamoeba castellanii*. These two lower eukaryotic systems provide an interesting and varied environment for studying the Thg1 family. *D. discoideum* encodes one Thg1 enzyme (DdiThg1)

with sequence similar to yeast and human Thg1, as well as three Thg1 like proteins (DdiTLP2, DdiTLP3, and DdiTLP4) (Figure 1A and 1B). In *A. castellanii*, however, there are only two Thg1-like proteins (AcaTLP1 and AcaTLP2) encoded in the genome, and no genuine Thg1 homolog. For each of the Thg1 superfamily enzymes in these two organisms, we have conducted a detailed biochemical study of multiple activities with respect to several types of tRNA processing events; we present the results of those studies here.

Chapter 1: Evolution of Thg1 enzymes in *Dictyostelium discoideum*: Specialization of 3'-to-5' polymerization for tRNA editing.

In most organisms, tRNA genes (tDNA) are transcribed as precursor tRNA (pre-tRNA) and go through extensive maturation steps including 5' processing by RNase P, 3' processing by various endo- and exonucleases, including RNase Z, and, in some cases, CCA adding enzyme and splicing to remove intronic sequences. In addition, all known tRNA species are subject to addition of various nucleotide base and sugar modifications. One uncommon alternative to tRNA biogenesis observed in *Nanoarchaeum equitans* is the formation of tRNAs from two separate genes encoding a 3' half and a 5' half (Randau et al., 2005). Here, we study yet another pathway for tRNA biogenesis that utilizes 3'-to-5' polymerization, a novel tRNA editing pathway found in the mitochondria of various lower eukaryotes including *D. discoideum*, *A. castellanii*, and *P. polycephalum*. In these organisms, mitochondrial tRNA genes encode mismatches at the first three positions of the acceptor stem, and if conventionally transcribed and processed, would lead to mitochondrial tRNAs (mt-tRNAs) with unconventional structure (Figure 3). Previous studies have shown that in *A. castellanii* and *P. polycephalum,* the encoded mismatches from the 5'-end are not maintained in the mature tRNA, and are replaced by nucleotides that restore normal base pairing between the 5'- and 3'-ends of the tRNA (Gott et al., 2010; Price and Gray, 1999a). Recently, we demonstrated that organisms that require mitochondrial tRNA editing also genomically encode TLPs, and proposed that the *in vitro* 3'-to-5' addition activity of TLPs plays a role in restoring acceptor stem base pairing (Abad et al., 2011).

The mt-tRNA editing pathway is predicted to involve two steps. First, the nascent pre-tRNA is processed by an as of yet unidentified endonuclease (possibly RNase P), to remove the 5' mismatches. The truncated tRNA intermediate that results can then be recognized by the repair

function of TLPs to produce full-length tRNAs. Interestingly, the predicted mismatches in all mitochondrial tRNA species, regardless of organism, occur at positions 1-3. As such, our analysis focuses on tRNAs missing up to three nucleotides as substrates for DdiTLPs.

The four Thg1 homologs in *D. discoideum* provide a unique opportunity to characterize multiple roles acquired by Thg1 family members in one species. It was previously observed that DdiThg1 and DdiTLP2 are highly specific for $tRNA^{His}$ (Abad et al., 2011), and their roles within the organism is thought to be distinct. DdiThg1 is predicted to function in cytoplasmic $tRNA^{His}$ maturation and DdiTLP2, encoded with a mitochondrial targeting peptide, is believed to be the mitochondrial counterpart of the $tRNA^{His}$ maturation enzyme (Long and Jackman, unpublished). Here, we focus on the characterization of DdiTLP3 and DdiTLP4 as candidates in mitochondrial tRNA editing. The goals of this study were to identify the substrate specificity and the unique mechanisms gained by these enzymes to adapt 3'-to-5' polymerization for efficient tRNA repair. Involvement of DdiTLPs in the editing mechanism predictably requires indiscriminate recognition of all potential editing substrates, as well as addition of the correct number and identity of nucleotides to the 5'-truncated tRNA to restore a full-length tRNA. Previous studies of DdiTLP3 and DdiTLP4 focused on activity with two truncated mt-tRNA substrates (mt-$tRNA^{Ile}$ and mt-$tRNA^{Leu}$), and both enzymes were demonstrated to catalyze repair reactions that coincide with the activity required in the predicted tRNA editing pathway (Abad et al., 2011). Furthermore, DdiTLP3 is encoded with a mitochondrial targeting peptide (Figure 1B) which would support a role for this enzyme in mitochondrial tRNA editing, although the possibility of an unidentified targeting signal for DdiTLP4 to the mitochondrion cannot be excluded. Therefore, the question of whether one or both of these enzymes participate in the in vivo editing reaction remains unaddressed. In this case, a comprehensive study on the substrate specificity of

DdiTLPs is required, since the editing machinery must recognize a large number (8 of the 18) mt-tRNAs in *D. discoideum*, and the possibility that the two enzymes coexist to repair non-overlapping sets of mt-tRNA substrates should be addressed. Here we investigate the substrate specificity and mechanistic properties of DdiTLP3 and DdiTLP4.

Results

Optimizing the DdiTLP reaction. Previously, in vitro reactions involving TLPs were performed using buffer conditions that were developed to assay G_{-1} addition activity catalyzed by yThg1. Here, we describe alternative conditions that have improved the observed rates for DdiTLP catalysis by ~ 5 fold as compared to previous conditions. Initial reaction rates under steady-state conditions were determined at varied concentrations of NaCl and KCl, and in the presence of varied concentrations of divalent cations (Mg^{2+} and Mn^{2+}). Other components of the reaction buffer were not changed from the originally developed Thg1 enzyme buffer (25 mM HEPES pH 7.5, 0.2 mg/mL BSA, and 3 mM DTT). The optimum salt concentration was found to be 50 mM NaCl, and 10 mM Mn^{2+} was found to be a better divalent cation for DdiTLPs than Mg^{2+} (Figure 11). This result is somewhat surprising, since previous experiments with yThg1 demonstrated no dependence of reaction rate on the identity of the divalent cation in the reaction (Mg^{2+} vs Mn^{2+}) (Smith and Jackman, unpublished). Most reactions described in this study use the optimum conditions described. However, the steady state kinetic measurements were conducted using Thg1 assay buffer (Table 1).

Substrate specificity of tRNA editing TLPs. The mismatches in the mitochondrial tRNAs of *D. discoideum* are found in 8 of the 18 mt-tRNAs, and Thg1 enzymes participating in the proposed editing pathway must recognize all predicted editing substrates. We therefore examined the

activity of DdiThg1/DdiTLPs on 7 of the 8 predicted tRNA intermediates (only one of two isoforms of $tRNA^{Ile}$ with nearly identical sequences was used) in the editing pathway using a phosphatase protection assay. Briefly, 5'- ^{32}P monophosphorylated tRNAs transcribed with deletions ranging from 1-3 nucleotides at the 5'-end (specified in Table 1**)**, and NTPs corresponding to the appropriate templated nucleotide addition reactions were included in the DdiThg1/TLP reaction. Catalysis of nucleotide addition resulted in the protection of the 5'-labeled phosphate from dephosphorylation after phosphatase treatment. The product mixture was then digested by RNase A or T1 to produce oligonucleotide products that can be resolved from inorganic phosphate, indicating unreacted substrate, on thin-layer chromatography (TLC) (Figure 4A). In some cases, the oligonucleotides that result from addition of 5'-nucleotides are difficult to conclusively identify by TLC due to their variable lengths and nucleotide compositions; in these cases, products from the 3'-to-5' addition reactions were further verified by primer extension. Consistent with previous observations that DdiThg1 and DdiTLP2 are G_{-1}-adding enzymes specific for $tRNA^{His}$, neither enzyme recognized the truncated tRNAs tested here (Table 1**)**. However, DdiTLP3 and DdiTLP4 catalyzed robust addition of missing nucleotides to all 7 tested tRNAs, and both enzymes catalyzed nucleotide addition with similar efficiency, as judged by amounts of repaired tRNA product produced using similar enzyme concentrations. Furthermore, DdiTLP3 and DdiTLP4 recognized and repaired two truncated tRNAs in the *in vitro* activity assays that do not contain encoded mismatches *in vivo*, mt-$tRNA^{AlaUGC}$ Δ3 and mt-$tRNA^{CysGCA}$ Δ1, (Table 1), suggesting that both enzymes are nonspecific and recognize truncated tRNAs regardless of identity.

A representative assay is shown in Figure 2**,** where product spots Cp*CpG and UpCp*CpG corresponding to addition of the missing C_{+2} and C_{+2}/U_{+1} nucleotides, respectively, to

tRNA^{GlnUUG} Δ2 were resolved from unreacted tRNA, seen as the inorganic phosphate spot (Figure 4B). The indicated identities of the product species were confirmed by additional nuclease digestion and nucleotide dropout experiments. Surprisingly, other spots were observed to accumulate to high levels in the DdiTLP4 reaction, and primer extension analysis was conducted to identify these products. Briefly, a DNA primer complementary to the tRNA D – loop sequence was annealed and extended using reverse transcriptase (RT) to produce cDNA with variable lengths corresponding to the number of nucleotides added to the 5'-end of the tRNA by DdiTLP3-4. The RT products were then separated on a denaturing polyacrylamide sequencing gel. In the presence of all 4 NTPs, DdiTLP3 predominantly repaired the truncated tRNA to the expected full-length (Figure 4C lane 15). A lower amount of an additional band was also observed; whether this band corresponds to U_{-1} addition or to the propensity of reverse transcriptase to add one additional dNTP to the 3'-end of its reaction products cannot be distinguished. However, DdiTLP4 clearly catalyzed excess nucleotide addition reactions, producing aberrant tRNAs with up to 3 extra nucleotides past the expected stop at position +1 (Figure 4C lane 16). Polymerization past position +1 has previously been observed with yThg1, and the implications of this activity will be discussed below.

DdiTLP4 catalyzes the repair of mt-tRNA^{HisGUG} Δ4. The tRNA 5'-editing substrates identified so far all contain three or fewer mismatched nucleotides, suggesting that the 5'-repair pathway might have a limitation to only adding 3 missing nucleotides to a tRNA substrate. To determine whether this is a limitation on tRNA truncations recognized by DdiTLPs, we tested activity using the phosphatase protection assay on monophosphorylated mt-ptRNA^{HisGUG} missing four nucleotides from the 5' end. Surprisingly, only DdiTLP4 catalyzed efficient repair, whereas DdiTLP3 had no substantial activity (Figure 5A**).** Because of the size of the expected product

(GpApApGp*GpU) the spots seen could not be clearly verified using TLC. Therefore, primer extension was conducted. For the primer extension, it should be noted that the bands observed in lanes 2-12 (Figure 5B) at the +4 position (indicating a Δ3 tRNA) are a result of one extra addition by reverse transcriptase, and decreasing the ddNTP concentration from 100 μM to 5 μM in the RT reaction reveals the correct band indicating a Δ4 tRNA (band at the + 5 position, Figure 5B lanes 10 and 12). Similar to the activity observed with mt-tRNA^{Gln} Δ2, DdiTLP4 consistently added up to 3 nucleotides past the +1 position (Figure 5B lanes 7-10). Surprisingly, when provided with a 5' triphosphorylated mt-ppptRNA^{His} Δ4, which does not require adenylylation, both DdiTLP3 and DdiTLP4 catalyzed nucleotidyl-transfer and primer extension indicated that the majority of DdiTLP3 catalyzed reaction products were full-length, with a RT stop at + 1 (Figure 5B lanes 5 and 8). It is conceivable that with mt-tRNA^{His} Δ4, DdiTLP3 is limited at the activation step, since nucleotidyl-transfer can still occur.

Templated vs non-templated nucleotide addition. Another feature that distinguishes the 5'-editing reaction for tRNAs is the need for repair to occur with high fidelity in order to restore fully Watson-Crick base paired aminoacyl-acceptor stems. It has been shown previously that DdiTLP3 and DdiTLP4 prefer to catalyze templated nucleotide addition, making base pairs with 3' end of the tRNA (Abad et al., 2011). Here, we sequenced the 5' ends of repair products of DdiTLP4 reacted with the mt-tRNA^{His} Δ4 substrate described above. To determine tRNA sequence, 4 separate reverse transcriptase reactions were carried out, each using a different ddNTP, with the ratio of ddNTP: dNTP as 1: 2. Incorporation of ddNTPs during the elongation of the cDNA results in chain termination, the length of which correlates with the sequence of the 5' tRNA end and can be identified on a sequence grade polyacrylamide gel. Consistent with the expected need for high-fidelity 3'-to-5' polymerase activity, sequencing revealed a complete set

of fully base paired nucleotides starting from position 4 (G_{+4}: C_{69}, A_{+3}: U_{70}, A_{+2}: U_{71}, and G_{+1}: C_{72}) corresponding to templated addition across the 3' tRNA end (Figure 5C**)**.

Though templated nucleotide addition is the preferred activity of TLPs, DdiTLP3 and DdiTLP4 were capable of non-templated nucleotide addition. Reactions were carried out using $ptRNA^{Gln}$ Δ2 (requiring C_{+2} and U_{+1}) and either DdITLP3 or DdiTLP4 with a combination of 1 mM CTP or UTP (with 0.1 mM ATP required for activation). Reactions performed in the presence of only 1 mM CTP and 0.1 mM ATP (missing the correct Watson-Crick pairing U_{+1} nucleotide) resulted in nucleotide addition up to position +1 for DdiTLP3 and -1 for DdiTLP4, which corresponds to the addition of C_{+2}: G_{71} and C_{+1}: A_{72} (and C_{-1}: A_{73} for DdiTPL4) (Figure 4C lanes 5-6). Results were similar with 1 mM UTP and 0.1 mM ATP, but nucleotide addition extended to the -1 position, corresponding to the addition of U_{+2}: G_{71}, U_{+1}: A_{72}, and U_{-1}: A_{73}, with possible U_{-2}: A_{74} addition by DdiTLP4 (Figure 4C lanes 7-8). This suggests that although DdiTLPs prefer to catalyze templated nucleotide addition in the presence of all NTPs, they can also catalyze non-templated nucleotide addition much like the non-templated G_{-1} addition activity seen with eukaryal Thg1.

Rates of polymerization by DdiTLP4. Due to the unexpected polymerization products catalyzed by DdiTLP4, a detailed study of the mechanism was required. To elucidate possible rate differences between truncated repair and excessive polymerization, the steady state kinetic parameters of polymerization by DdiTLP4 at various tRNA truncations was measured using 5' labeled p*pptRNAs in an assay that measures release of the labeled pyrophosphate from the 5'-end of the tRNA upon nucleotide addition to the activated 5'-end. The optimal reaction conditions for DdiTLP 3'-to-5' polymerization had not been identified when this kinetic study was initiated, and therefore, all kinetic measurements reported were conducted with yThg1

activity buffer. *In vitro* transcription of substrates with triphosphorylated 5' ends required transcripts to initiate with a guanosine. We therefore chose ppptRNA^{LeuUAG} from *A. castellanii* in the following study, which contains 4 guanosine residues at the first four nucleotide positions, and synthesized a set of ppptRNA^{LeuUAG} substrates, each missing from 1-3 nucleotides at the 5' end. Reactions were conducted with 1 mM GTP and tRNA under steady state conditions, where the tRNA concentration was 5 fold or greater than the concentration of DdiTLP4. The initial reaction rates (up to 10% product conversion) were normalized for enzyme concentration, and $V_o/[E_o]$ at various tRNA concentrations was fitted to the Michaelis-Menten equation to calculate k_{cat}, K_M, and k_{cat}/K_M. The kinetic parameters for polymerization to Δ1, Δ2, and Δ3 ppptRNA^{Leu} were similar, with a k_{cat}/K_M around 1 x 10^5 $M^{-1}s^{-1}$. Interestingly, there was a kinetic preference for addition to truncated tRNAs, with a significant decrease (~ 230 fold) in kinetic efficiency for nucleotide addition on full-length substrates (Table 2). Furthermore, the kinetic measurements conducted here, on the ppptRNA^{Leu} substrate, show ~ 470 fold increase in kinetic efficiency as compared to a monophosphorylated p-tRNA^{Ile}, consistent with the increased kinetic efficiency of addition to ppp-tRNA over p-tRNA observed for other Thg1 family enzymes (Abad et al., 2011; Jackman and Phizicky, 2006a).

ATP inhibition of DdiTLP4 activity. The effects of various nucleotides on the nucleotidyl-transfer step were tested using a pyrophosphate release assay under conditions similar to those in the phosphatase protection assay. p*pp tRNA^{LeuUAG} Δ3 and combinations of GTP + NTP (where N = C, U, or A) were assayed with DdiTLP4, and incorporation of nucleotides resulted in the release of labeled inorganic pyrophosphate (PPi) that could be resolved from unreacted tRNA (p*ppptRNA) on TLC. Inclusion of combinations of GTP + CTP and GTP + UTP in the assay caused a modest decrease of ~ 1.2 fold in specific activity as compared to a GTP only reaction

(Figure 6**)**. The presence of ATP, however, had a drastic effect on nucleotidyl-transfer, with ~ 12 fold reduction in specific activity, and ADP had an intermediate effect, with a ~ 4 fold decrease in activity. A non-hydrolysable analog of ATP, α – thio ATP, was also tested, and the effect on activity was similar to ATP, suggesting that ATP consumption (such as its use for activation of 5'-ends) was not required for this inhibitory effect.

To understand the effect of ATP on nucleotidyl-transfer at different positions of truncation, we repeated measurements of the steady state kinetic parameters on mt-ppptRNA^{Leu} substrates previously tested (Table 2). We observed a significant decrease (from 25 fold to 110 fold) in the overall kinetic efficiency of nucleotide addition to the truncated tRNAs in the presence of ATP, with k_{cat}/K_M aound 1 x 10^3 $M^{-1}s^{-1}$. Moreover, the effect of ATP on k_{cat} and K_M varied with position, where the decreased kinetic efficiency for addition to Δ3 and Δ2 ppptRNA^{Leu} was a result of a decreased in the rate constant only. The change in K_M on these substrates was within the error of the reactions containing only GTP. However, the decrease in k_{cat}/K_M on Δ1 ppptRNA^{Leu} was due to a ~50 fold decrease in K_M, and ATP only had a modest ~2 fold decrease in k_{cat}. More importantly, the inhibiting effect of ATP extends to the -1 position, where the rate of nucleotide addition to full-length ppptRNA^{Leu} in the presence of ATP was decreased by ~82 fold. The k_{cat}/K_M for addition to full-length tRNA is 6.8 $M^{-1}s^{-1}$, about 440 fold slower than the average rates for addition on truncated substrates tested in the presence of ATP. These data suggest a role for ATP-mediated control of the repair reaction catalyzed by DdiTLP4.

Primer extension can be used to identify the 5' length of tRNAs. However, to better understand the effect of ATP on all steps of DdiTLP4-catlayzed repair (including the activation step), a more readily quantifiable method was required. We therefore chose 5' labeled p*tRNA^{Leu} Δ3 encoded in the mitochondria of *D. discoideum* as it produces a 6-nucleotide fragment after RNase T1

digestion which can be resolved from longer repair products on a 15% polyacrylamide gel. Both DdiTLP3 and DdiTLP4 were assayed with this method in the presence of all 4 NTPs, and as previously seen by primer extension, DdiTLP3 stops after adding 3 nucleotides to produce full-length tRNA and DdiTLP4 adds an additional nucleotide at the -1 position (Figure 7A). The experiment was repeated, only this time ATP was titrated to observe the effect of ATP on N-1 addition by DdiTLP4 (Figure 7B). Increasing ATP concentrations, from 1 mM to 6 mM, slightly increased the total amount of product produced from 43 % to 46 %, and concentrations above 6 mM caused complete inhibition. Interestingly from 1 mM to 6 mM ATP, the fraction of N-1 decreased drastically from 19 %, nearly half of the total product, to 5 %. For DdiTLP4, 6 mM ATP is an optimal concentration for its repair capabilities, reducing the amount of N_{-1} addition to 10 % of the total amount of product formed.

Discussion

Our focus on mitochondrial tRNA editing has led to the characterization of DdiTLP3 and DdiTLP4, identifying a broad substrate specificity and unique mechanisms to mediate 3'-to-5' polymerization for tRNA repair. Unlike their G_{-1} adding counterparts, DdiTLP3 and DdiTLP4 are indiscriminate with respect to tRNA identity, and repair all truncated tRNAs tested with similar efficiency. However, the results in this study indicated important differences between DdiTLP3 and DdiTLP4. First, detailed study of the products of tRNA repair catalyzed by DdiTLP3 (Figure 4C and 3B**)** suggests that the inherent ability to recognize tRNA length, since the majority of the tRNA products of DdiTLP3 catalyzed repair generate a full-length tRNA (with a 5'-terminal +1 nucleotide), regardless of the number of nucleotides that need to be added to each substrate. DdiTLP4, however, catalyzed a polymerization reaction that also exhibits similarities to what is observed with yThg1, where nucleotides were incorporated past the

expected full-length of the tRNA. It is predicted that pyrophosphate removal catalyzed by yThg1 (deactivating the 5' end from further nucleotide additions) is a mechanism to ensure a single G_{-1} addition to $tRNA^{his}$ Interestingly, DdiTLP4 also is subject to a mechanism to prevent these predictably "excess" nucleotide additions. In this mechanism, ATP reduces the kinetic efficiency of polymerization reaction at all positions, including truncated positions, decreasing the total amount of 5'-extended tRNA produced in each reaction (Table 2). However, the kinetic efficiency of the repair reaction with a pre-activated ppp-$tRNA^{Leu}$ in the presence of 1 mM ATP is ~ 9 fold faster than with a p-$tRNA^{Ile}$ substrate (Table 2), suggesting the rate limiting step in the kinetic scheme of DdiTLP4 catalysis may be activation. Initial results (not shown) support this conclusion, since a 5 - 6 fold difference between the observed rates for activation and nucleotidyl-transfer have been measured. The presence of up to 6 mM ATP during the repair of truncated $ptRNA^{Leu}$ reduced the amount of N_{-1} formed while maintaining near constant product conversion (Figure 7B). More importantly, the range of ATP concentrations maximizing the amount of full-length product formed (4-6 mM ATP) is reasonable, given the physiological concentrations of ATP in a typical cell. *In vivo,* cellular concentrations of ATP could allow controlled nucleotide addition by DdiTLP4, preventing polymerization past the full-length tRNA end.

Due to a lack of structural information and a detailed understanding of the mechanism of TLP catalysis, it is unclear how ATP inhibits DdiTLP4 activity. Using the previously published crystal structure of hThg1, along with the data presented, we have established a testable model to describe ATP inhibition. The identified active site of hThg1 suggests the involvement of two divalent metal ions, and it is believed that the three carboxylates conserved among all Thg1 family members help in coordinating the metal ions for catalysis (Figure 1B) (Hyde et al., 2010;

Smith and Jackman, 2012). From the hThg1 crystal structure, two nucleotide-binding sites are observed, one for ATP used in adenylylation and the other for GTP used in nucleotidyl-transfer. Due to conservation of residues around the active sites among DdiTLPs and hThg1 (Abad et al., 2011) and similar 3'-to-5' addition activity, it is conceivable that TLPs share similar active site architecture to hThg1. The ATP-specific inhibition observed in our study suggests DdiTLP binds specifically to ATP during polymerization. We predict that ATP inhibition occurs when ATP binds to the ATP binding site normally used for adenylylation, since after the first activation and nucleotide addition, the resulting ppptRNA does not require ATP for subsequent additions. (Figure 11). There are several reasons as to why ATP binding may inhibit DdiTLP4. Due to the close proximity of the 5' triphosphorylated tRNA end, ATP, and the incoming nucleotide at the active site, the binding of ATP may reduce the affinity or sterically interfere with tRNA or NTP binding, or even both. Furthermore, it is predicted that Thg1 proteins catalyze nucleotide addition via a two metal ion active site, which would require correct coordination of the metal ions to favor catalysis. It is conceivable that the triphosphate of ATP alters the conformation of the metal ions in a position that is less optimal for catalysis, leading to a decrease in the rate of nucleotidyl transfer. The kinetic measurement of DdiTLP4 on different truncation lengths of $tRNA^{leu}$ suggests the effect of ATP varies with position, where for $tRNA^{leu}$ $\Delta 2$ and $\Delta 3$, ATP decreased the k_{cat} while having no significant effect on K_M. For $tRNA^{leu}$ $\Delta 1$, ATP increased the K_M while mildly decreasing the k_{cat}. It is possible, therefore, that the mechanism of ATP inhibition varies with tRNA length.

Evolutionarily, DdiTLP4 and Thg1 enzymes evolved separate mechanisms to control 3'-to-5' polymerization, where the former uses ATP to regulate addition and the latter catalyzes removal of inorganic pyrophosphate (PPi). It has been previously predicted that the ancestral activity of

Thg1 enzymes may have been general 3'-to-5' RNA polymerization (Jackman and Phizicky, 2006b). It is conceivable, therefore, that both ATP inhibition and PPi removal have been acquired through time as 3'-to-5' polymerization was set aside for tRNA biogenesis. Furthermore, the polymerization reactions catalyzed by TLPs may be more closely related to the ancestral activity than G_{-1} addition, since the majority of TLPs are found within Bacteria and Archaea whereas Thg1 proteins are only found in Eukarya. As such, Thg1 may have evolved as a specialization of TLPs for $tRNA^{His}$ maturation. The capability of DdiTLP4 to catalyze N_{-1} addition (Figure 4C, 5B, and 7A), as well as non-templated nucleotide addition (Figure 4C) suggests TLPs are flexible enzymes, and through selective pressure, may have acquired the capability to add G_{-1} to $tRNA^{His}$. Previous characterization of the TLP in *Bacillus thuringiensis* (BtTLP) described the ability to catalyze tRNA repair and add G_{-1} to $tRNA^{His}$, serving as a link between Thg1 and TLP activities (Rao et al., 2011b). Therefore G_{-1} addition may have arisen from evolutionary pressures to retain $tRNA^{His}$ identity, and a TLP similar to DdiTLP4 or BtTLP may have been the ancestor to eukaryal Thg1.

Interestingly, the flexibility of DdiTLP4, with the ability to catalyze repair of $tRNA^{His}$ Δ4 as well as polymerization past the +1 position, contrasts with the very limited window of tRNA repair catalyzed by DdiTLP3. With the exception of $tRNA^{His}$ Δ4, DdiTLP3 catalyzed repair up to the full-length position (Figures 2, 3 and 5) on all the truncated tRNAs tested. It is reasonable to conclude that DdiTLP3 has evolved intrinsic mechanisms to identify tRNA length, compared to the ATP assisted mechanism used by DdiTLP4. Combined with evidence that tRNA mismatches occur only at the first three nucleotide positions and DdiTLP encodes a mitochondrial targeting peptide, the limited repair activity of DdiTLP3 and tRNA editing in the mitochondria of *D.*

discoideum may have evolved together, and over time, DdiTLP3 has become specialized in mt-tRNA editing.

However, it is unlikely that tRNA editing existed before the introduction of TLPs into the mitochondria. Possibly, the presence of TLPs in the mitochondria of an ancestral organism allowed for the accumulation of mismatches at the 5' end of mt-tRNAs. This would also require the presence of a nuclease that could recognize tRNA mismatches, yet the nuclease involved in tRNA editing has not been identified. New evidence on the subcellular localization of DdiTLP3 and DdiTLP4 indicate distinct compartmentalization of these enzymes. Green-fluorescent-protein (GFP) fused DdiTLP3 is localized in mitochondria only (Long and Jackman, unpublished), while DdiTLP4 is localized in the cytosol. It is unclear whether DdiTLP4 enters the mitochondria, and since it also lacks a mitochondrial targeting peptide, the longer repair capabilities of DdiTLP4 may therefore indicate a role as a general cytosolic tRNA repair enzyme. This would entertain the possibility that all four DdiTLPs have been specialized in both activity and locality, where in the cytosol DdiThg1 and DdiTLP4 function in G_{-1} addition and tRNA repair, respectively, and in the mitochondria DdiTLP2 and DdiTLP3 (both containing mitochondrial targeting peptides) function in G_{-1} addition and tRNA editing, respectively.

Chapter 2: Is G_{-1} on $tRNA^{His}$ truly ubiquitous in Eukarya? A study of the evolutionary relationship between the Thg1 superfamily and the G_{-1} requirement.

To understand the evolutionary connection between Thg1 enzymes and tRNA maturation, it is important to understand the various pathways Thg1 family members have taken to function in tRNA biology. The study of DdiTLPs provided an avenue in understanding how the presence of 3'-to-5' tRNA repair can influence tRNA genes and possibly allow alternative pathways for tRNA biogenesis. Here, we investigate the integrity of the relationships established by Thg1 enzymes and tRNAs, and how in eukaryotes, G_{-1} on $tRNA^{His}$ can be associated as closely with Thg1 as it can be with HisRS.

As discussed above, G_{-1} on $tRNA^{His}$ was discovered to be the primary identity element for HisRS, and in eukaryotes, it is believed that G_{-1} is added post–transcriptionally across A_{73} by eukaryal members of the Thg1 superfamily. However, several eukaryotes have been identified that do not encode a true Thg1 ortholog (Jackman et al., 2012). We focus our study on the properties of the Thg1 homologues in *A. castellanii*, a protozoan encoding two TLPs (AcaTLP1 and AcaTLP2), and the roles they may play in tRNA editing and in $tRNA^{His}$ maturation. *A. castellanii* TLPs fall under the bacterial and archaeal subgroup and have sequences similar to DdiTLP3 and DdiTLP4. Like DdiTLP3, AcaTLP2 contains an N-terminal mitochondrial targeting peptide (Figure 1B) and we predict they play a role in tRNA repair and mitochondrial tRNA editing, which has been extensively analyzed in the past in *A. castellanii* (Burger et al., 1995; Lonergan and Gray, 1993; Price and Gray, 1999a, b). Interestingly, the cytoplasmic $tRNA^{His}$ species in *A. castellanii* contains an A_{73} discriminator nucleotide, and the -1 residue of the nascent pre-tRNA transcript is a uridine. Notably, it is unlikely that the AcaTLPs catalyze G_{-1} addition to the cytoplasmic $tRNA^{His}$, considering all TLPs previously characterized required a

C_{73} containing $tRNA^{His}$ for G_{-1} addition. As a result, G_{-1} cannot be obtained genomically, and the status of the $tRNA^{His}$ encoded in *A. castellanii* remains a mystery. We therefore investigated the activity of AcaTLPs to understand the relationship between the Thg1 superfamily and $tRNA^{His}$ maturation within the organism.

Results

AcaTLPs catalyze truncated tRNA repair, not G_{-1} addition. To understand the role AcaTLPs play in tRNA maturation within *A. castellanii*, the activities of purified AcaTLP1, AcaTLP2, as well as a AcaTLP2 construct lacking the mitochondrial targeting peptide (AcaTLP2-mito) were tested on several of the truncated tRNAs prepared for the study of DdiTLPs, including a native *A.castellanii* $tRNA^{Leu}$ Δ2, using a phosphatase protection assay (Figure 4A). Like their *D. discoideum* counterparts, AcaTLPs catalyzed robust tRNA repair on all truncated tRNAs tested, and interestingly, both enzymes recognized and repaired the $tRNA^{His}$ Δ4 transcript (a feature which distinguishes DdiTLP3 and DdiTLP4) (Table 1).

The lack of a Thg1 ortholog in *A. castellanii* brings into question whether AcaTLPs can catalyze non-templated G_{-1} addition on $tRNA^{His}$. Using a phosphatase protection assay, we tested for G_{-1} addition on an A_{73} containing $ptRNA^{His}$ from *S, cerevisiae*. Small amounts of App*GpC were seen, indicating activation of the tRNA, but neither of the AcaTLPs catalyzed G_{-1} addition (indicated as Gp*GpC) (Figure 8A). On the surface, this might be considered difficult to rationalize, since *A. castellanii* would predictably require post-transcriptional G_{-1} addition, as is observed in other eukaryotes. We then tested G_{-1} addition activity on C_{73} containing mt-$tRNA^{His}$ from *D. discoideum* (Figure 8B). AcaTLP2 with and without the mitochondrial-targeting peptide catalyze G_{-1} addition, as indicated by the Gp*GpU spot produced after RNase digestion of the

products. Other spots were observed on the TLC, which may indicate longer polymerization products. AcaTLP1, however, has no activity, and therefore its only function may be tRNA repair. These results suggest that AcaTLP2 may play a role in mitochondrial G_1 addition.

Yeast complementation assay. To further support our *in vitro* analysis of AcaTLPs, we conducted a yeast complementation assay to test whether AcaTLPs could complement the G_{-1} addition activity of yThg1. Briefly, AcaTLPs were cloned downstream of a galactose-inducable promoter in a yeast overexpression plasmid containing a *LEU2* marker. The plasmid was then transformed into a yeast strain (Δthg1, Δleu2, Δura3) that contained a plasmid carrying the yeast *THG1* under its native promoter and a URA3 marker. The transformed strain was then grown on minimal selective media (SGal – leu), and then replica plated onto media containing 5-fluorourotic acid (SGal – Leu + 5FOA), which selects against the *URA3/THG1* containing plasmid while promoting the expression of the AcaTLPs being tested. Growth on media containing 5FOA indicates complementation of the yeast Thg1 activity, and AcaTLP - containing strains did not support growth (Figure 9). This is likely due to a lack of G_{-1} activity, as our *in vitro* studies suggested.

Discussion

Our study of DdiTLPs has shown that the availability of 3'-to-5' tRNA repair may have affected tRNA evolution, allowing mismatches to accumulate in the mitochondrial tRNAs of various lower Eukarya. Now, we show that the absence Thg1 may indicate the absence of G_{-1} addition in *A. castellanii*. Like *D. discoideum*, *A. castellanii* requires tRNA editing to restore correct tRNA structure to several mt-tRNAs. We show here that both AcaTLP1 and AcaTLP2 catalyze robust repair of truncated tRNAs, and since AcaTLP2 is encoded with a mitochondrial-targeting

peptide, these TLPs may play a role in a possible tRNA editing pathway similar to that in *D. discoideum*. Furthermore, DdiTLP2 catalyzes efficient templated G_{-1} addition, suggesting it may partake in $tRNA^{His}$ maturation in the mitochondria. However, the lack of a Thg1 ortholog brings to question the status of the G_{-1} requirement on the cytoplasmic $tRNA^{His}$. Though bacterial species have been characterized to encode and retain G_{-1} during $tRNA^{His}$ maturation, pre-$tRNA^{His}$ in *A. castellanii* likely encodes a U_{-1}, and neither of the AcaTLPs catalyze non-templated G_{-1} addition.

The recently sequenced 5'-end of cytoplasmic $tRNA^{His}$ in *A. castellanii* (Rao and jackman, unpublished) strikingly showed an absence of G_{-1}, supporting the *in vitro* and *in vivo* data described here (Figure 10). The lack of G_{-1} brings into question the identity elements used by HisRS to recognize and histidylate $tRNA^{His}$. Previously, all eukaryal HisRSs studied required G_{-1}, with the exemption of a handful of alpha-proteobacteria that contained HisRS that recognizes the GUG anticodon of $tRNA^{His}$. Whether the loss of the G_{-1} requirement is due to recognition of the GUG anticodon by the *A. castellanii* HisRS, or to the establishment of new identity elements by the synthetase is unknown, and warrants further investigation of *A. castellanii* HisRS. Nonetheless, the lack of both G_{-1} and Thg1 in *A. castellanii* suggests the retention of Thg1 is correlated with the need for G_{-1} addition in eukaryotes. Surprisingly several other eukaryotes, including *C. elegans,* do not encode Thg1 enzymes yet are predicted to require G_{-1} addition activity. Therefore, phylogenetic analysis could potentially be used to identify organisms with unusual $tRNA^{His}$ identity elements.

This study on the activities catalyzed by Thg1 organisms describes the possible involvement of the Thg1 family in both mitochondrial tRNA and $tRNA^{His}$ maturation. 3'-to-5' polymerization by Thg1 enzymes have provided alternative pathways to produce tRNAs capable of participating in

translation by maintaining base pairing in mt-tRNAs of at least the two organisms described here. We observed that Thg1 enzymes such as DdiTLP4, have evolved mechanisms to allow correct repair of tRNAs, even though it is capable of excessive polymerization, Furthermore, we see that the interplay between Thg1, HisRS, and the G_{-1} requirement on $tRNA^{His}$ in eukaryotes can be used to identify organisms that lack the G_{-1} requirement.

Materials and Methods

Thg1/TLP expression and purification. AcaTLP1 and AcaTLP2 genes were PCR-amplified from *A. castellanii* cDNA (total RNA kindly provided by Dr. Michael Gray, Dalhousie University). Since the AcaTLP2 gene sequence contains a predicted mitochondrial-targeting peptide, constructs were created to produce proteins with and without this N-terminal mitochondrial-targeting peptide (AcaTLP2 ± mito). PCR products were cloned into AVA421 plasmid downstream of an N-terminal His_6 and an IPTG-inducible promoter for over-expression in *E. coli*. Constructs for expression and purification of DdiThg1/TLPs were previously described (Abad et al., 2011). Thg1/TLP expression constructs were transformed into *E. coli* strain Rosetta pLysS, and the desired proteins were over-expressed and purified by immobilized metal-ion affinity chromatography as described previously (Abad et al., 2010). All resulting proteins were dialyzed in 50% glycerol for storage and assessed for purity (≥ 90%) by SDS-PAGE. Protein concentrations were determined using the Bio-Rad protein assay. The concentrations of each purified protein ranged from 9-160 µM.

***In vitro* synthesis and 5' [^{32}P] labeling of tRNAs.** All truncated and full-length tRNA transcripts were synthesized using the T7 RNA polymerase in vitro transcription protocol described in (Abad et al., 2011). Briefly, tRNA constructs were synthesized using pairs of

overlapping oligonucleotides and fill-in polymerization, and ultimately cloned into a pUC19 vector using appropriate restriction sites found in the multiple cloning site. tRNAs initiating with a 5' guanosine were constructed directly downstream of a T7 RNA polymerase promoter, thus producing 5' triphosphorylated tRNAs (ppp-tRNAs), either labeled or unlabeled depending on the presence of radiolabeled nucleotides in the reactions. tRNAs initiating with 5'-A, C, or U were placed downstream of a T7 promoter and a hammerhead ribozyme sequence designed to produce the specified tRNAs with a 5'-hydroxylated end following in vitro transcription and co-transcriptional ribozyme cleavage, which was assisted by inclusion of a temperature cycling step (10 cycles of 60 ^{o}C x 3 min → 25 ^{o}C for 3 min) to the reactions. A BstNI restriction site was introduced into all constructs via PCR to produce the correct 3'-CCA after run-off transcription**.** *A. castellanii* mt-tRNA^{LeuUAG} Δ2 and Δ3 were constructed via Quikchange mutagenesis (Stratagene) of a tRNA^{Leu} Δ1 construct.

Transcription products were gel-purified using 10% polyacrylamide 4 M urea. To produce 5' [α-^{32}P] monophosphorylated tRNAs (p-tRNAs) , ppp-tRNA transcripts were treated with calf-intestinal phosphatase (NEB) to produce 5'-hydroxylated ends (not necessary for tRNAs transcribed with a hammerhead ribozyme leader) and subsequently treated with T4 polynucleotide kinase and [γ-^{32}P] ATP. Excess ATP was removed using BioGel P6 columns (BioRad) and another round of gel purification. To produce 5'[γ-^{32}P] triphosphorylated tRNAs (limited to tRNAs initiating with 5'-G), transcription reactions were carried out using [γ-^{32}P] GTP.

Optimization of reaction conditions. Previously, activity assays for DdiTLPs were conducted in Thg1 assay buffer containing 10 mM $MgCl_2$, 125 mM NaCl, 25 mM HEPES pH 7.5, 0.2 mg/mL BSA, 3 mM DTT, and a combination of NTPs depending on the tRNA assayed (Jackman

and Phizicky, 2006a). To identify optimal reaction conditions for DdiTLPs, the rate of ptRNA^{Gln} Δ2 repair by DdiTLP4 under steady state conditions (see steady state assay conditions) were compared after altering the salts (0-125 mM NaCl or 0-125 mM KCl) and divalent cations (0-10 mM $MgCl_2$ or 0-10 mM $MnCl_2$) used in the yeast Thg1 assay (Figure 12). The optimized DdiTLP assay buffer was determined to be 50 mM NaCl, 10 mM $MnCl_2$, 0.2 mg/mL BSA, 3 mM DTT, and 25 mM HEPES pH 7.5.

Phosphatase protection assay. To test nucleotide addition activity on truncated and full-length tRNAs, a phosphatase protection assay was used as described in (Jackman and Phizicky, 2006a). Briefly, 10-50 nM 5'-^{32}P-labeled ptRNAs were reacted with fivefold serial dilutions of purified enzyme (10 μM-1.4 nM) in either DdiTLP assay buffer or Thg1 assay buffer as indicated in the text for 1-2 h at room temperature. Reactions were quenched with EDTA, RNase digested (T1 or A) and subsequently phosphatase treated. Reaction products were resolved on silica TLC in a 55:35:10 (v:v:v. 1-propanol: NH_4OH; H_2O) solvent system. TLC plates were visualized using a Typhoon Trio phosphorimager. Reaction of mt-tRNA^{Leu} Δ3 (Figure 7) with DdiTLP3 and DdiTLP4 was alternatively processed. RNase T1 digestion of the reaction products produced RNA fragments 6-10 nucleotides long, which were resolved on 15 % polyacrylamide, 4 M urea.

Primer extension assay. mt-tRNA^{gln} Δ2 mt-tRNA^{HisGUG} Δ4 from *D. discoideum* were reacted with DdiTLP3 and DdiTLP4 and analyzed using primer extension. DdiTLP reactions contained 1-2 μM of ptRNA or ppptRNA, 0.9 μM of DdiTLP3 or 9.4 μM DdiTLP4, and a combination of NTPs as indicated in the text. Reactions were carried out for 1-2 h, and the resulting tRNAs were purified by phenol extraction and ethanol precipitation. 2-3 pmol of tRNA product were used as a template for primer extension with AMV reverse transcriptase and 1 pmol 5'- ^{32}P-labeled DNA

primer specific for the tRNA tested. Primer extension products were resolved on 10 % polyacrylamide, 4 M urea sequencing gels.

Steady state assay. Steady-state kinetic measurement of DdiTLP4 activity was measured using 5' [γ-^{32}P] mt-ppptRNA^{Leu} missing 1, 2, or 3 nucleotides (all G residues). The buffer used in the reactions was not optimized (10 mM $MgCl_2$, 125 mM NaCl, 25 mM HEPES pH 7.5, 0.2 mg/mL BSA, 3 mM DTT), and reactions were conducted with tRNA concentrations (0.08 - 12 μM) maintained at a minimum of 5 fold greater than DdiTLP concentrations (1 - 400 nM). TLP-catalyzed nucleotide addition to the labeled 5'-end causes the release of labeled inorganic pyrophosphate, which can be resolved from unreacted substrate using PEI-Cellulose TLC in a 80:20 0.5 M potassium phosphate pH 6.3:methanol system. Time courses measured the initial reaction rate, V_o (<10 % product formation), at various tRNA concentrations and normalized for the amount of enzyme used, V_o/[E]. Reaction rates were plotted as a function of tRNA concentration and kinetic parameters were derived after fitting to the Michaelis-Menten equation:

$$\frac{V_o}{[\mathrm{E}]_o} = \frac{k_{cat}[\mathrm{tRNA}]}{\mathrm{K_M} + [\mathrm{tRNA}]}$$

Yeast complementation assay. Thg1 and TLP genes were cloned downstream of a galactose-inducable promoter (P_{GAL}) into plasmids containing a *LEU2* marker and transformed into yeast strain JJY20 (relevant genotype: *thg1Δ, ura3Δ, leu2Δ*). The JJY20 strain also contains a *URA3* plasmid with wild-type yeast *THG1* under its native promoter (P_{THG1}). Positive transformants were selected on synthetic dextrose (SD) media lacking leucine and replica plated to synthetic

media containing galactose (SGal), to induce expression of the Thg1 and TLPs tested, and FOA, to select against the y*THG1*-containing *URA3* plasmid, thus leaving the heterologous THG1/TLP gene as the only source of Thg1 activity in the cells.

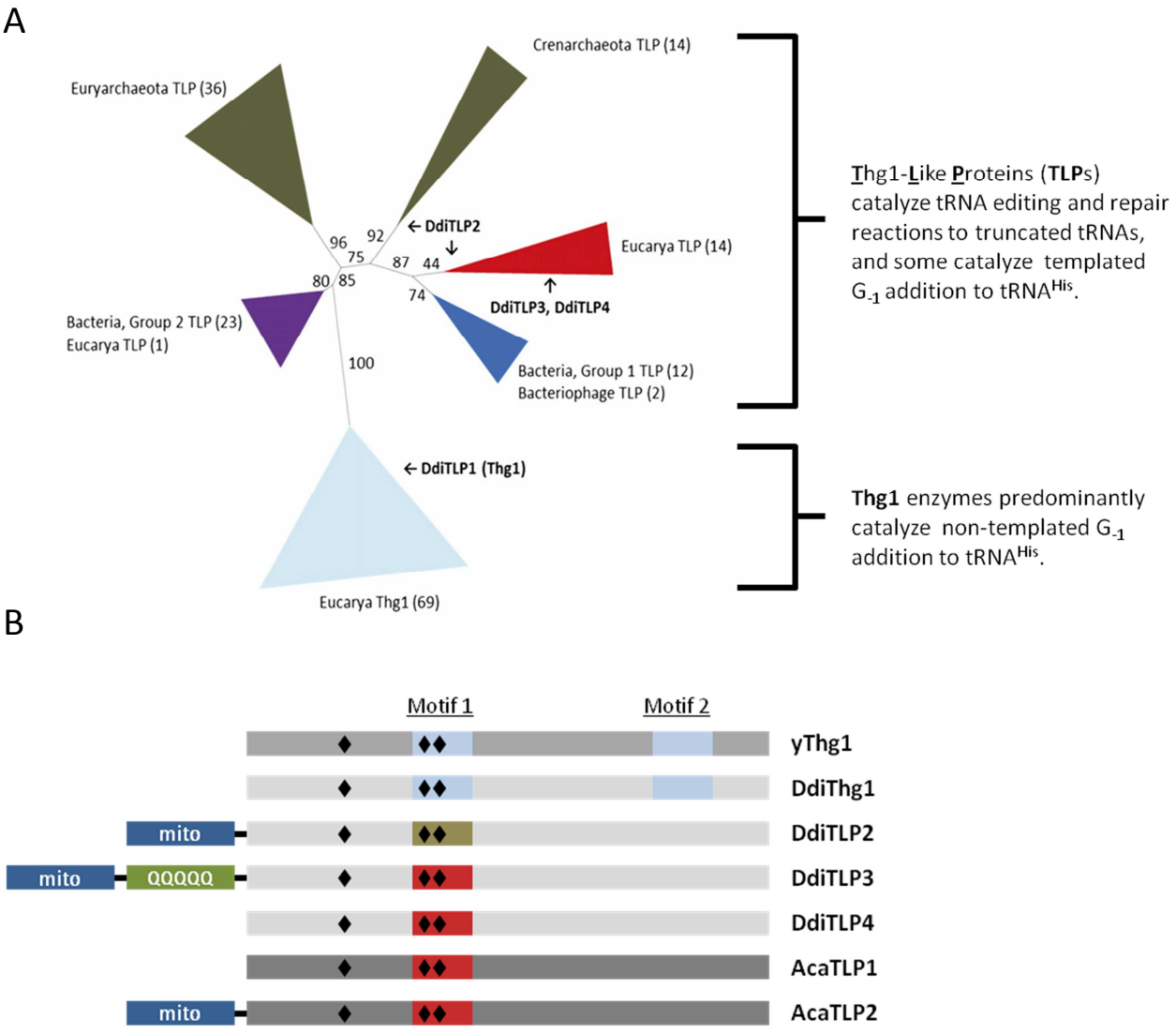


Figure 1: Phylogeny of the Thg1 Superfamily. (A) Phylogenetic analysis of the Thg1 superfamily suggests two distinct subgroups. Enzymes studied from the eukaryal Thg1 subgroup catalyze non-templated G_{-1} addition to $tRNA^{His}$, while enzymes from archaeal and bacterial subgroup catalyze tRNA repair, templated G_{-1} addition, or both. Numbers in parenthesis indicate the number of sequences used within the branch, and numbers at junctions correspond to bootstrap values. (B) Illustration of Thg1/TLP sequences show three conserved carboxylates (black diamonds), all of which are located in the proposed active site of the enzymes to coordinate metal ions necessary for catalysis. Various motifs are conserved within subgroups, where the eukaryal subgroup contains a conserved HINNLYN motif 2 and a motif 1 sequence distinct from that of TLPs. It should be noted that DdiTLP2, DdiTLP3, and AcaTLP2 all contain predicted mitochondrial targeting peptides.

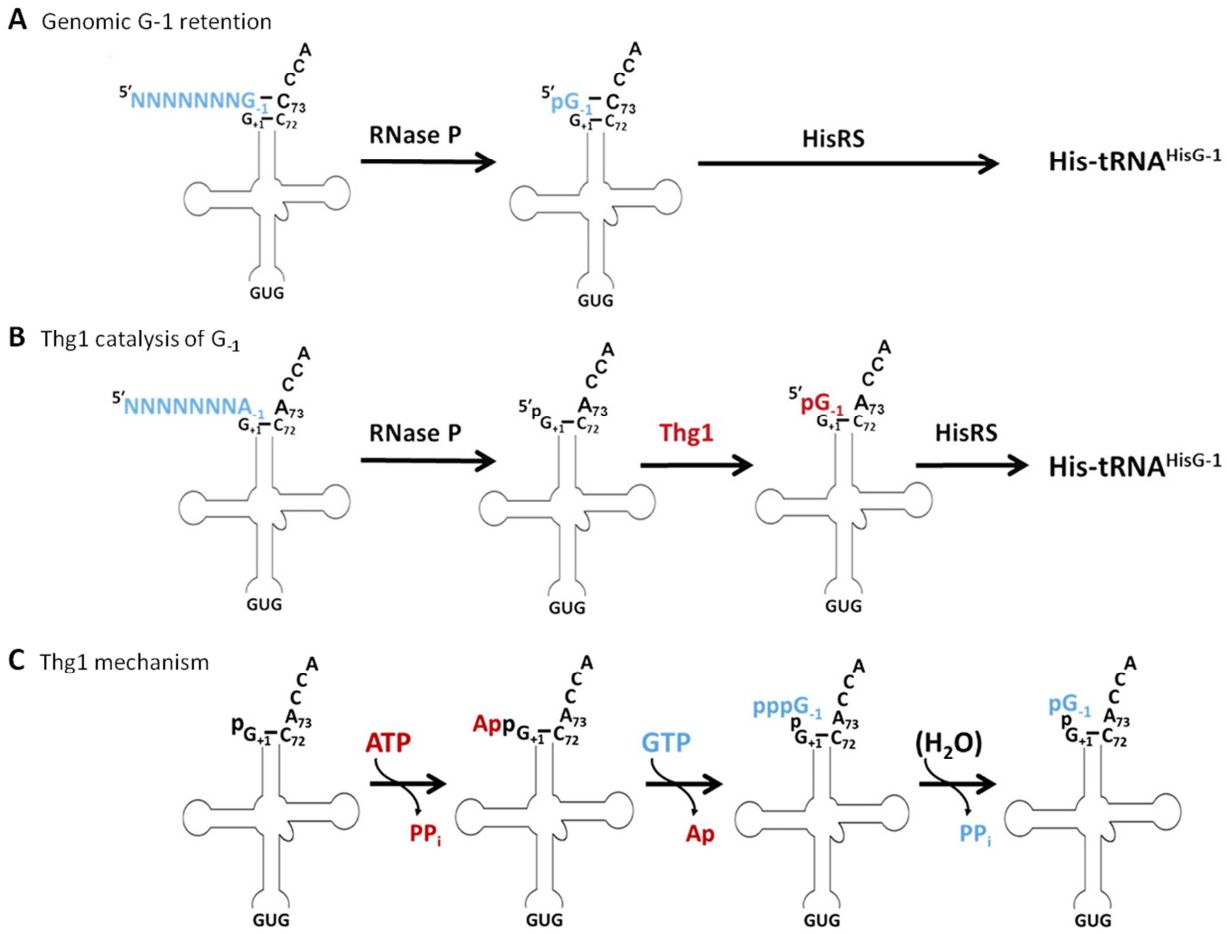


Figure 2: Mechanisms for G_{-1} incorporation and Thg1 catalysis. (A) In many bacterial and some archaeal species, G_{-1} is encoded in the tDNA and retained after 5' processing of $tRNA^{His}$. (B) All studied eukaryal $tRNA^{His}$ obtain the G_{-1} residue post-transcriptionally by Thg1. (C) The mechanism of G_{-1} addition follows three chemical steps: First, Thg1 catalyzes adenylylation of the 5' monophosphorylated end to activate for subsequent nucleophilic attack by GTP, adding the G_{-1} residue. Finally, Thg1 catalyzes the removal of pyrophosphate to yield G_{-1} containing $ptRNA^{HIs}$.

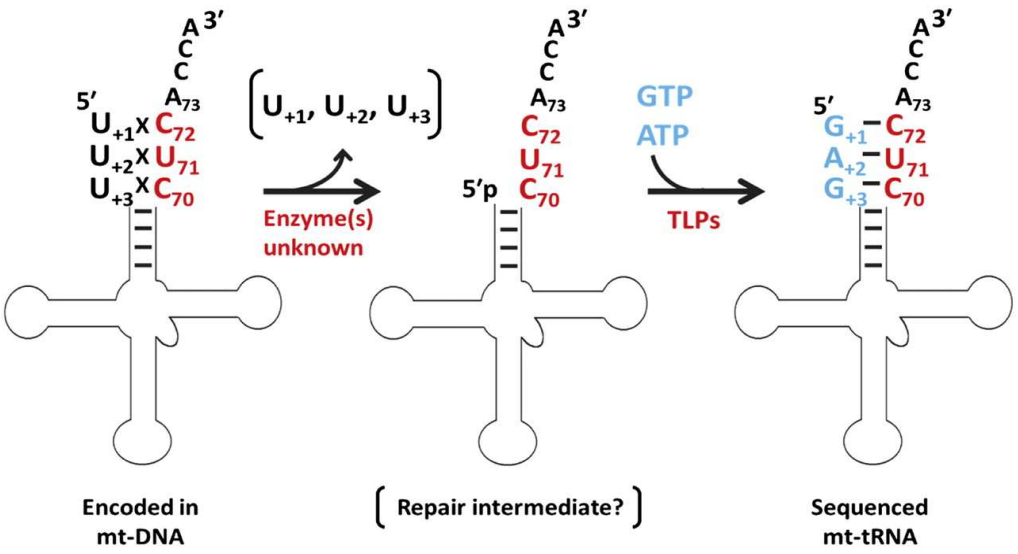
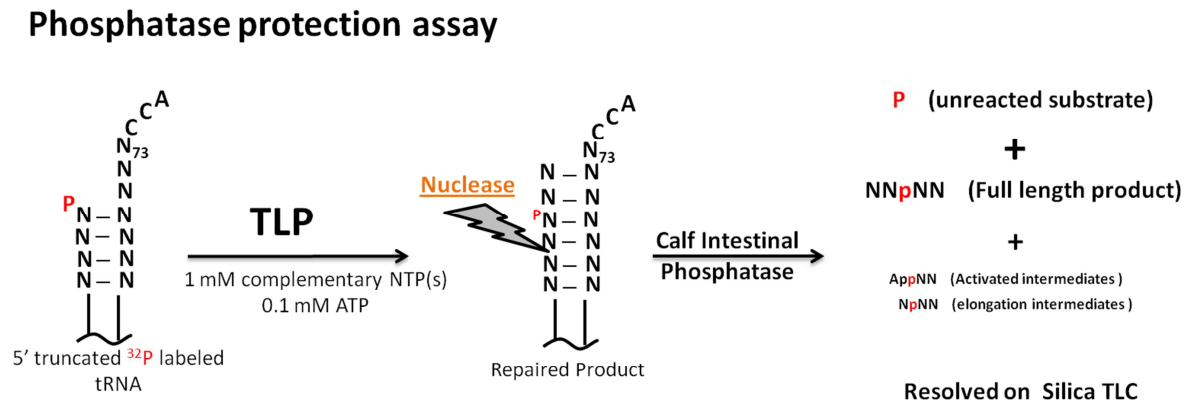
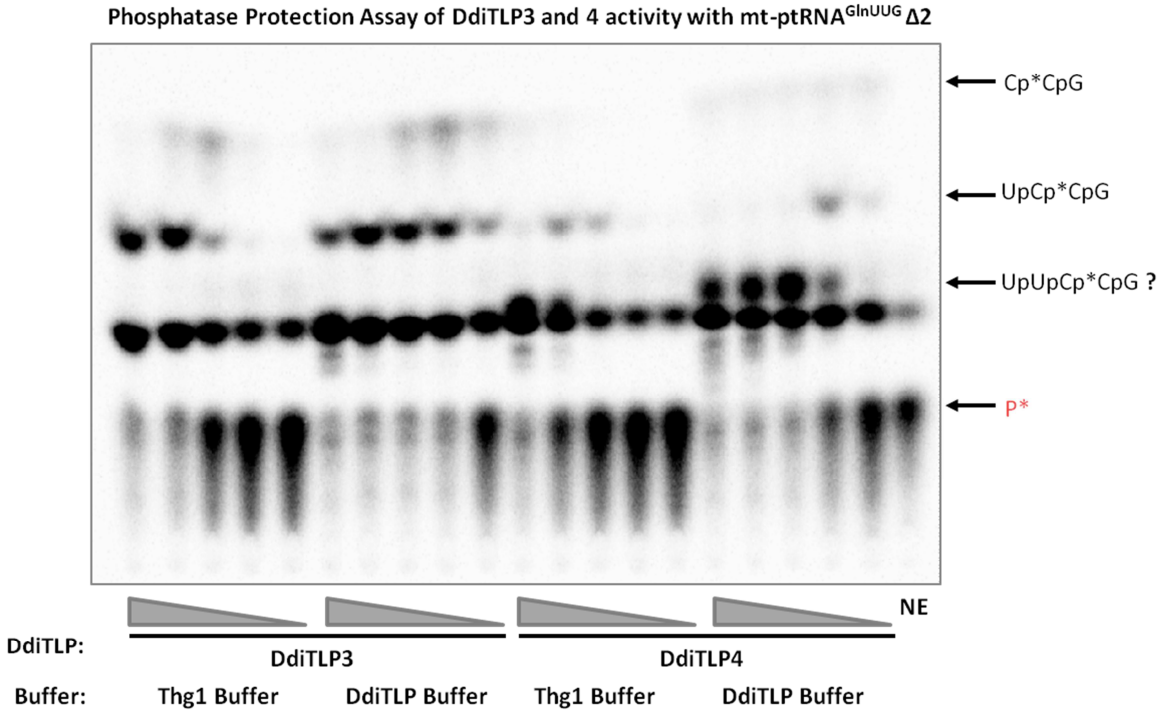


Figure 3: Hypothesized tRNA editing pathway The proposed editing pathway initiates with the cleavage of mismatched nucleotides encoded at the 5’ end of mitochondrial tRNAs by a(n) unknown enzyme(s). The resulting truncated tRNA intermediates can then be substrates for templated 3’-to-5’ tRNA repair by TLPs to produce a full length tRNA.

A


Phosphatase protection assay
TLP
1 mM complementary NTP(s)
0.1 mM ATP
Nuclease
Calf Intestinal
Phosphatase
5' truncated 32P labeled
tRNA
Repaired Product
P (unreacted substrate)
+
NNpNN (Full length product)
+
AppNN (Activated intermediates)
NpNN (elongation intermediates)
Resolved on Silica TLC

B


Phosphatase Protection Assay of DdiTLP3 and 4 activity with mt-ptRNA GlnUUG Δ2
Cp*CpG
UpCp*CpG
UpUpCp*CpG ?
P*
NE
DdiTLP:
DdiTLP3
DdiTLP4
Buffer:
Thg1 Buffer
DdiTLP Buffer
Thg1 Buffer
DdiTLP Buffer

C

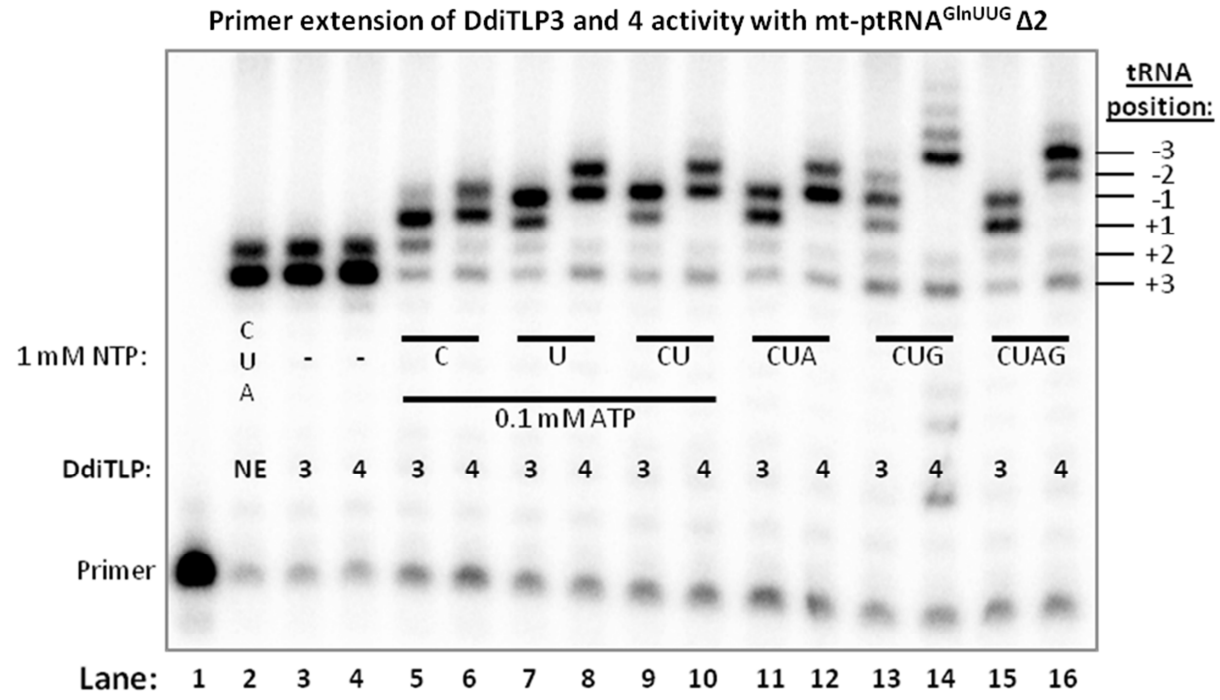


Figure 4: Representative activity assay with DdiTLP4. (A) The schematic for the phosphatase protection assay is shown. Briefly, incorporation of nucleotides to the 5'-^{32}P monophosphorylated tRNA (p*tRNA) by TLPs protects the labeled phosphate from phosphatase activity. The repair products are then digested by either RNase A or T1, and resolved on silica TLC. (B) A representative TLC scan of DdiTLP activity on mt-ptRNA^{Gln} Δ2 after RNase T1 digestion. Spots correlating to C_{+2} addition (Cp*CpG) and U_{+1} addition (UpCp*CpG) are indicated. DdiTLP4 reactions have an additional spot that migrates lower, presumably indicating U_{-1} addition (UpUpCp*CpG). (C) Primer extension of the repair of mt-ptRNA^{Gln} Δ2. The nucleotides used in the DdiTLP reactions are indicated, with lanes 15 and 16 showing DdiTLP activity with all nucleotides present.

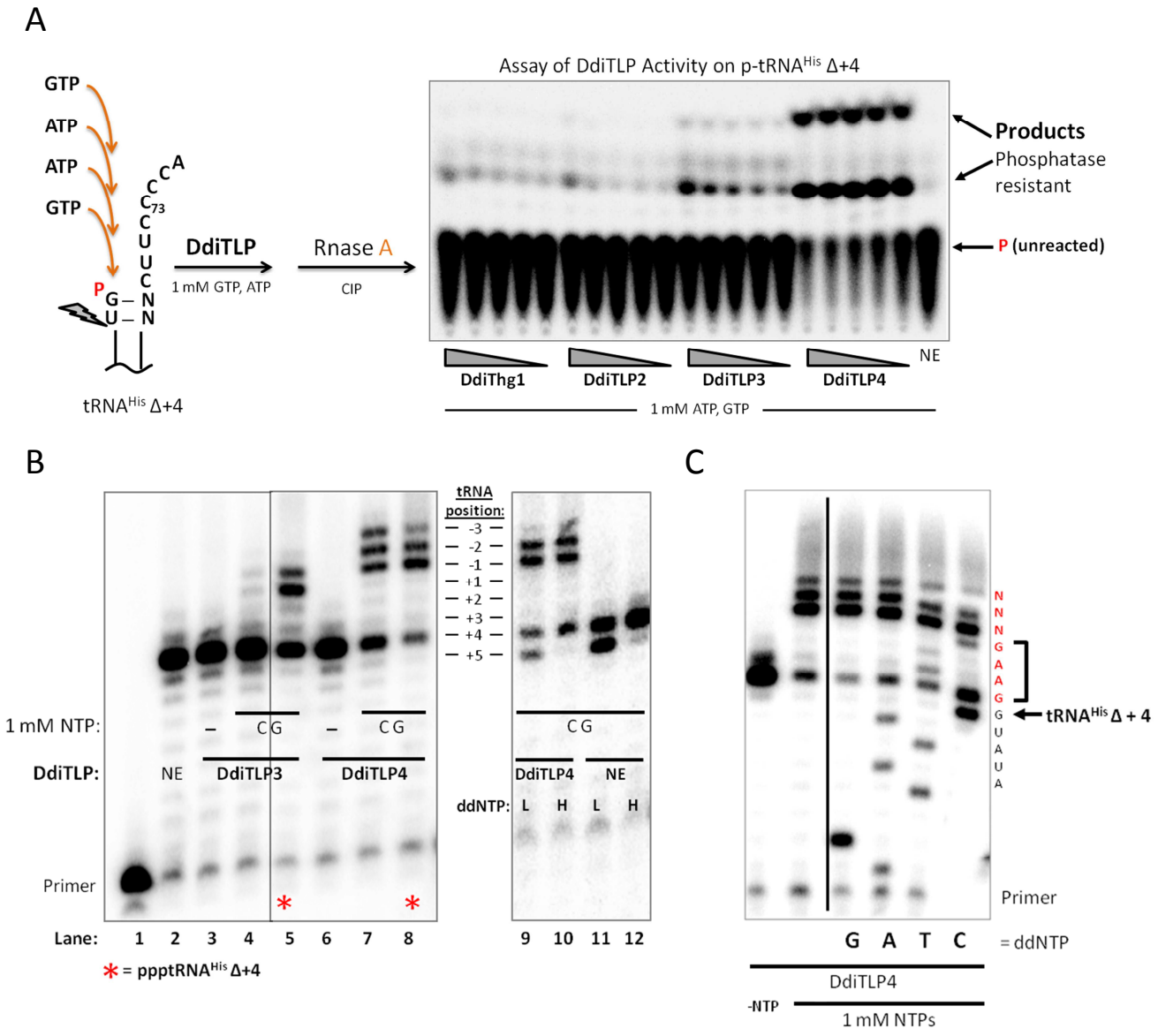


Figure 5: Primer extension analysis of tRNA^{His} Δ4 repair. (A) Phosphatase protection assay on mt-ptRNA^{His} Δ4. DdiTLP4 alone displays significant repair activity, though the identity of the product spots could not be established using TLC. (B) Primer extension of the repair products of DdiTLPs with mt-tRNA^{His} Δ4. Bands at the +4 position (the tRNA is expected to start at the +5 position) in lanes 2-12 indicate one extra dNTP addition by RT, which was alleviated by decreasing the dNTP concentration from **H = 100 µM to L = 5 µM** in lanes 9-12. Triphosphorylated tRNAs were used in lanes 5 and 8. (C) The repair of mt-ptRNA^{His} Δ4 in the presence of all NTPs and DdiTLP4 was sequenced by primer extension. The sequence in red indicates the nucleotides incorporated by DdiTLP4.

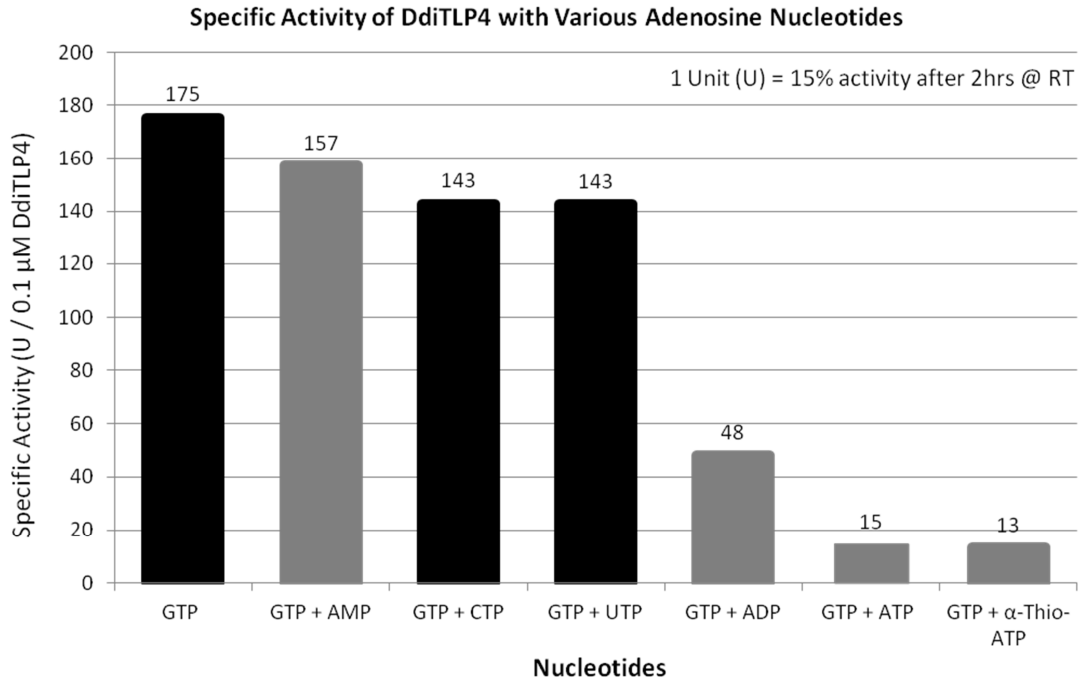
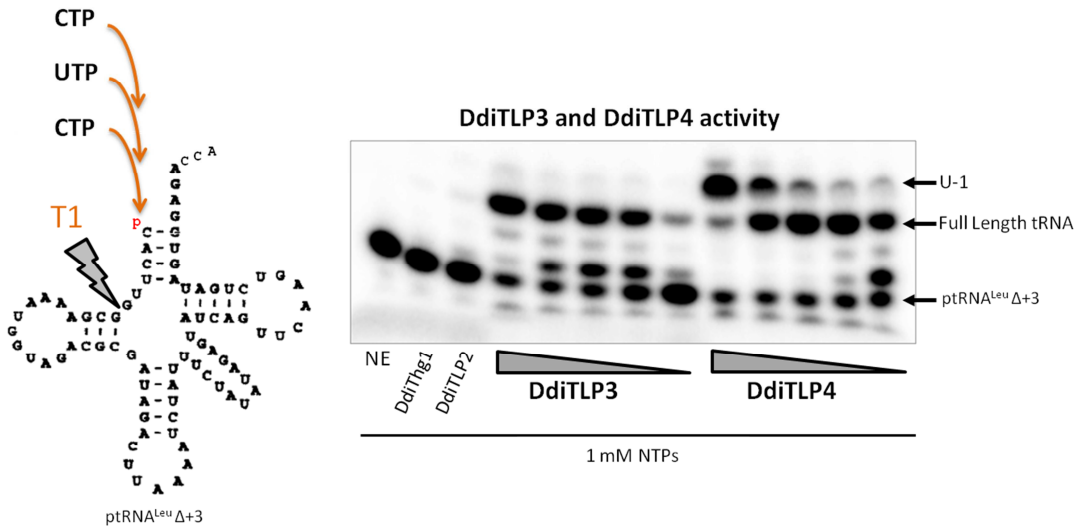


Figure 6: The affect of different nucleotides on DdiTLP4 repair. The repair of triphosphoryated mt-tRNA^{leu} Δ3 required the addition of $G_{+3,}$ G_{+2}, and G_{+1}. The effect of other nucleotides during GTP incorporation by DdiLTP4 was measured. Adenosine analogues are highlighted in grey.

A



B

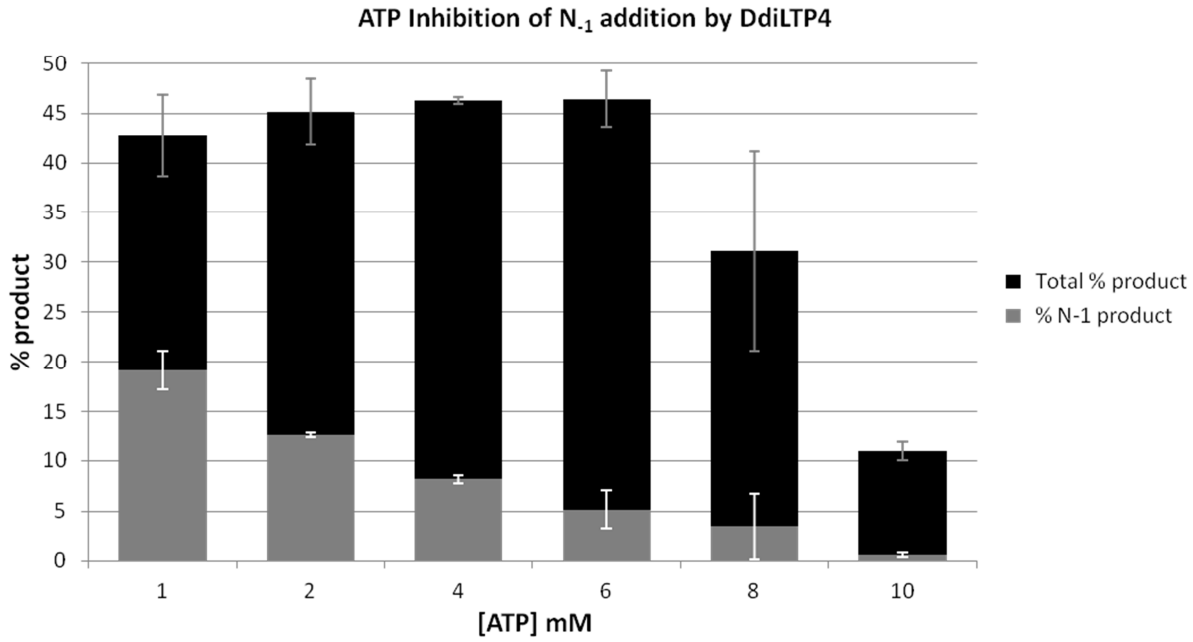
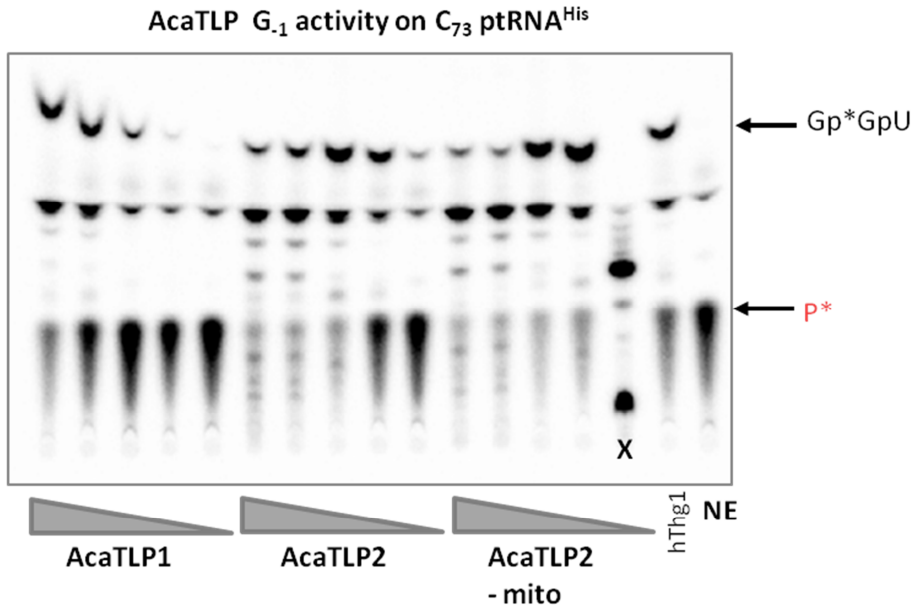


Figure 7: ATP inhibition of DdiTLP4. (A) To quantify the effect of ATP inhibition, monophosphorylated mt-tRNA^{Leu} Δ4 (missing C_{+3}, U_{+2}, and C_{+1}) was tested with the DdiTLPs. Digestion with T1 produces a 6 – 10 nt RNA fragment that was resolved on 15 % polyacrylamide, 4M urea. Bands correlating to full length repair and U_{-1} addition are indicated. (B) The assay was repeated with DdiTLP4 and varying ATP concentrations (1-10 mM as indicated),. The amount of U_{-1} addition was compared to the total amount of product formed. 6 mM ATP maximizes the amount of full length tRNA formed while minimizing U_{-1} addition.

A



B

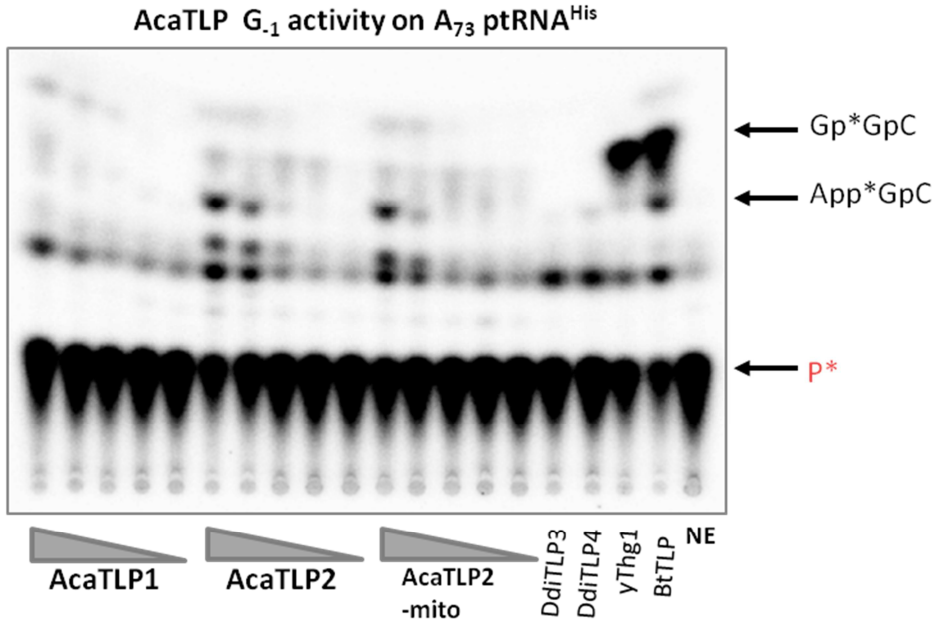


Figure 8: Activity of AcaTLPs on tRNA^{His}. (A) AcaTLP2 catalyzes templated G_{-1} addition to C_{73} tRNA^{His} (Gp*GpC spot), and possibly polymerizes, while AcaTLP2 has much weaker activity. (B) Neither of the AcaTLPs catalyzes non-templated G_{-1} addition to A_{73} tRNA^{His}.

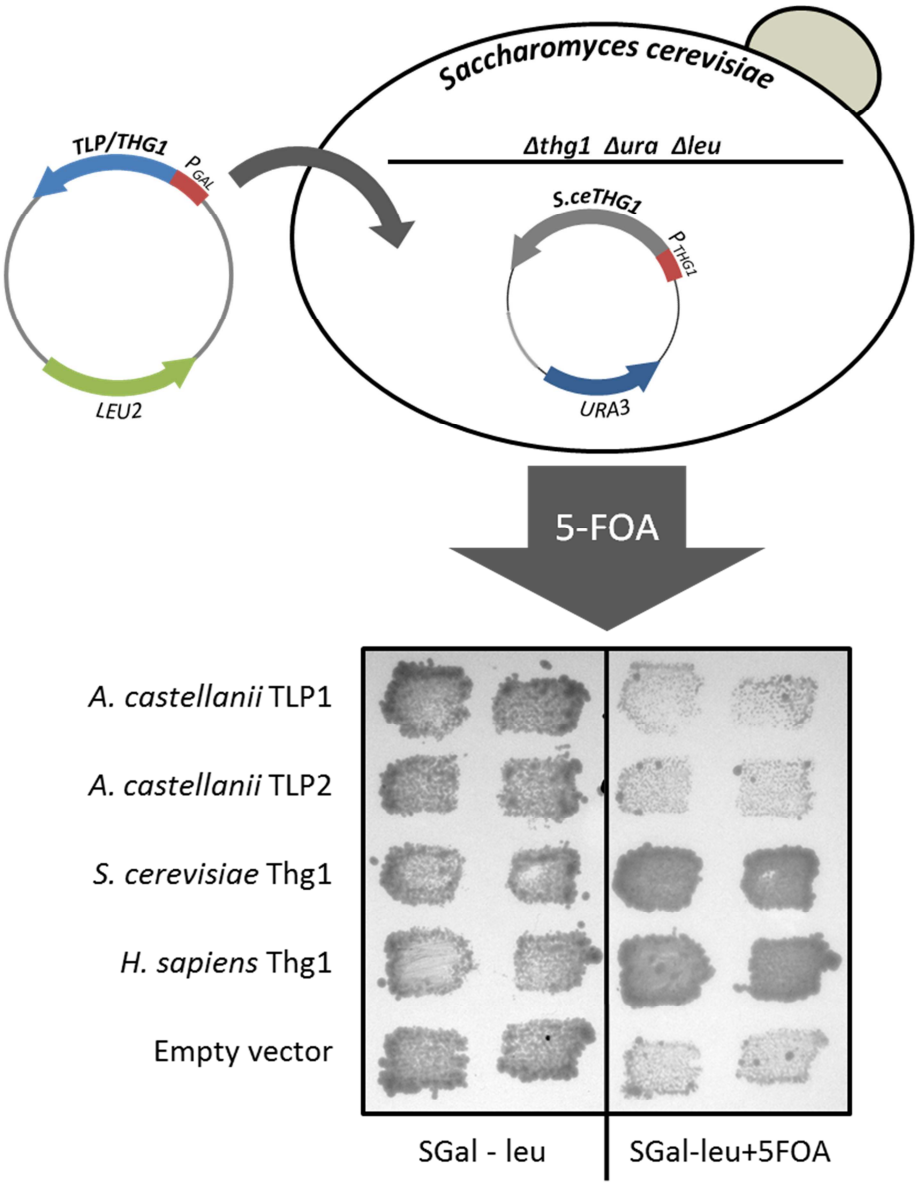


Figure 9: Yeast Complementation Assay. Neither AcaTLP1 nor AcaTLP2 complement the G_{-1} addition activity of yThg1 in a yeast Δthg1 strain. A *LEU2* plasmid with the various Thg1/TLPs indicated was transformed into a yeast Δthg1 Δura Δleu strain supplemented with yThg1 on a *URA3* plasmid. Strains were replicated on the media described and grown for 3 days at 30 ^{o}C.

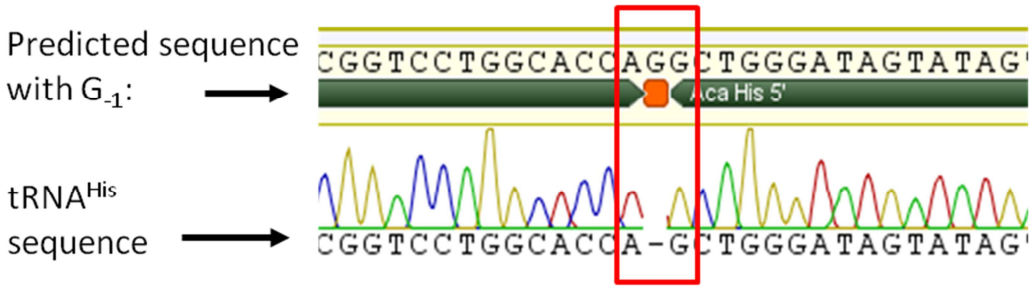


Figure 10: Sequence of *A. castellanii* cytoplasmic tRNA^{His}. Cytoplasmic tRNA^{His} from total *A. castellanii* RNA was circularized and cloned by RT-PCR. The 11 sequence reactions conducted showed the same results, indicated above, where cytoplasmic tRNA^{His} lacks a G_{-1} residue. Data provided by Bhalchandra Rao, The Ohio State University.

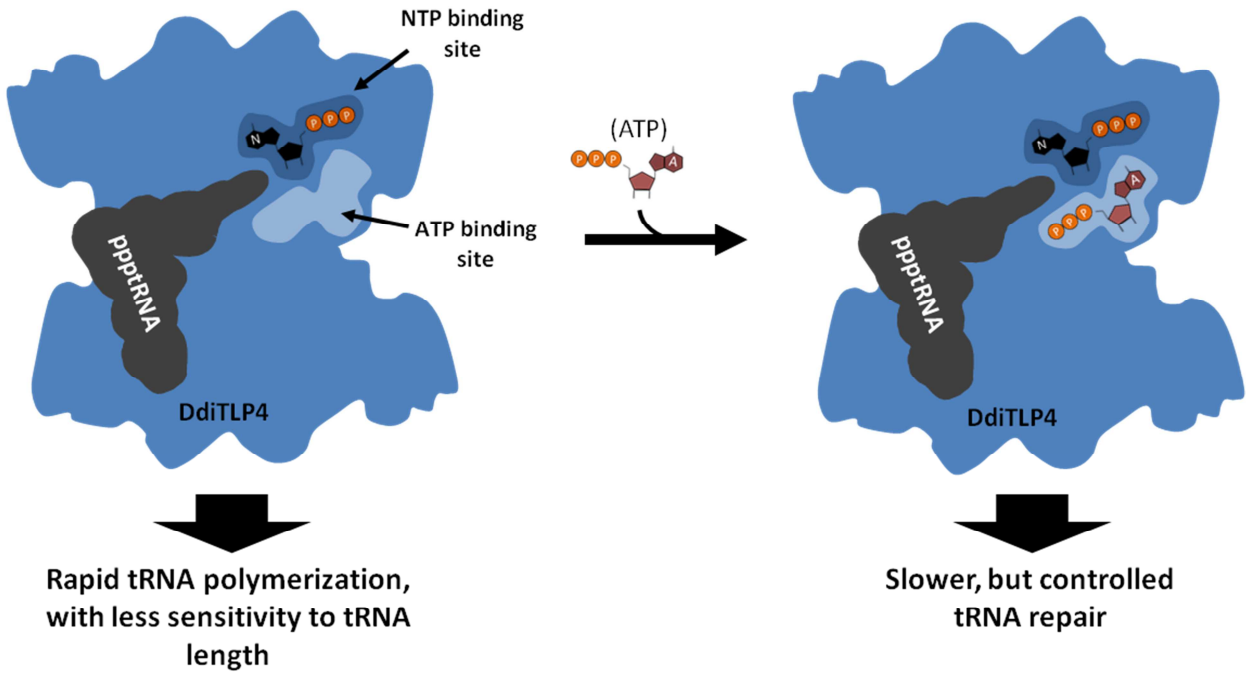


Figure 11: Illustration of the proposed mechanism for ATP inhibition of DdiTLP4. The active site of yThg1, and likely DdiTLP4, contains an ATP binding site for adenylylation, as well as a general NTP binding site for nucleotidyl transfer. We predict that ATP bound to DdiTLP4 during the nucleotidyl-transfer step reduces the rate of polymerization, allowing for better tRNA length recognition.

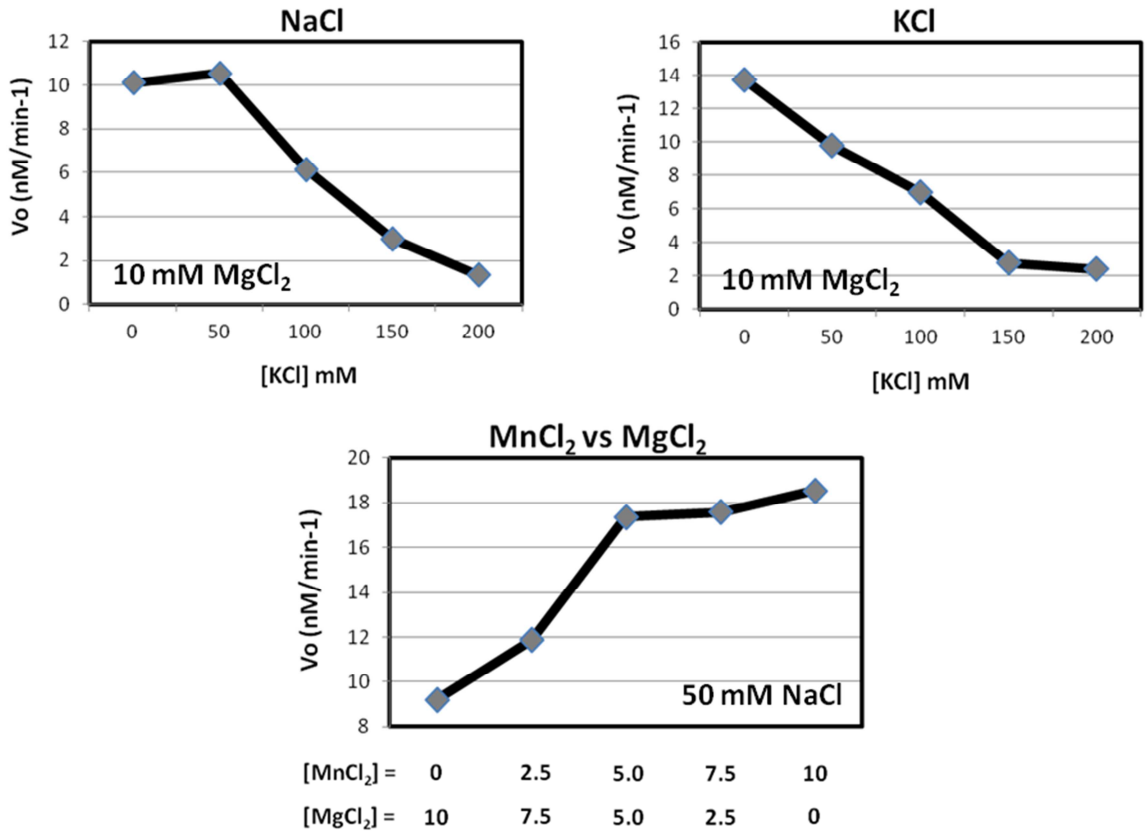
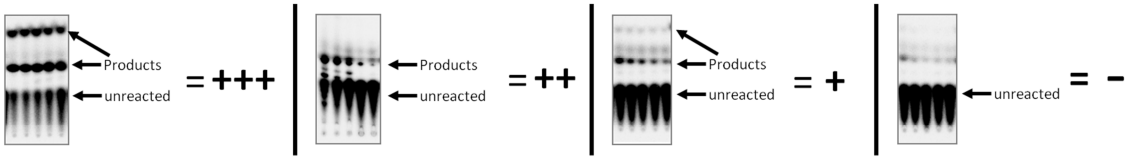


Figure 12: Optimization of DdiTLP activity buffer. Different salt conditions and divalent metal ions were tested once to optimize the activity buffer for DdiTLPs. The following condition was found to be optimum: 50 mM NaCl, 10 mM $MnCl_2$, 0.2 mg/mL BSA, 3 mM DTT, and 25 mM HEPES pH 7.5.

Table 1: Summary of TLP activity on tested substrates

tRNA Substrates (From *D. discoideum unless stated)*	Truncation	Requires Editing?	DdiThg1	DdiTLP2	DdiTLP3	DdiTLP4	AcaTLP1	AcaTLP2
Polymerization across 3' CCA (any tRNA)	-	-	yes	no	no	- ATP		
S. cerevisiae His_{GUG} C_{73}	Full length	G_{-1}- C_{73}	++	+	-	-		
S. cerevisiae His_{GUG} A_{73}	Full length	G_{-1}- A_{73}	++	-	-	-	-	-
His_{GUG}	Full length	G_{-1}- C_{73}	+	+++	++	+	+	++
Cys_{GCA}	Δ+1	no	-	-	+++	+++		
Glu_{UUC}	Δ+1	Yes	-	-	+++	+++		
Ile_{CAU}	Δ+1	Yes	-	-	+++	+++		
Tyr_{GUA}	Δ+1	Yes	-	-	+++	+++		
A. castellanii Leu_{UAG}	Δ+2	Yes	-	-	+++	+++	+++	+++
Leu_{UAG}	Δ+2	Yes	-	-	+++	+++		
Ile_{GAU}	Δ+3	Yes	-	-	+++	+++		
Leu_{UAA}	Δ+3	Yes	-	-	+++	+++		
Gln_{UUG}	Δ+3	Yes	-	-	++	++	+++	+++
Ala_{UGC}	Δ+3	no	-	-	+++	+++		
His_{GUG}	Δ+4	G_{-1}- C_{73}	-	-	-	+++	+++	+++

Key:



Key comparing the relative activities measured by the phosphatase protection assay. Products resolved by Silica TLC

Table 2: Steady-state kinetic parameters of DdiTLP4 catalyzed repair

$ppptRNA^{Leu}$	1 mM NTP	1 mM ATP	k_{cat} (min^{-1})	$K_{M\ tRNA}$ (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
Full Length	UTP	-	0.059 ± 0.005	1.8 ± 0.6	560 ± 190
Δ1	GTP	-	1.84 ± 0.18	≤ 0.13	≥ 240000
Δ2	GTP	-	1.13 ± 0.05	0.15 ± 0.03	130000 ± 26000
Δ3	GTP	-	1.15 ± 0.06	≤ 0.20	≥ 96000
Full Length	UTP	+	0.0061 ± 0.0053	~ 15	6.8
Δ1	GTP	+	0.85 ± 0.08	6.6 ± 1.3	2100 ± 470
Δ2	GTP	+	0.079 ± 0.003	0.25 ± 0.05	5400 ± 1100
Δ3	GTP	+	0.079 ± 0.007	0.90 ± 0.24	1500 ± 410
Δ1	-	+	0.14 ± 0.08	~ 4.5	500
$ptRNA^{Ile}$	1 mM NTP	ATP	k_{cat} (min^{-1})	$K_{M\ tRNA}$ (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
Δ1	GTP	0.1 mM	0.05 ± 0.01	2.6 ± 1.4	330 ± 120

Note: all kinetic parameters were measured under sub-optimal yThg1 assay buffer.

References

Abad, M.G., Long, Y., Willcox, A., Gott, J.M., Gray, M.W., and Jackman, J.E. (2011). A role for tRNA(His) guanylyltransferase (Thg1)-like proteins from Dictyostelium discoideum in mitochondrial 5'-tRNA editing. Rna *17*, 613-623.

Abad, M.G., Rao, B.S., and Jackman, J.E. (2010). Template-dependent 3'-5' nucleotide addition is a shared feature of tRNAHis guanylyltransferase enzymes from multiple domains of life. Proceedings of the National Academy of Sciences *107*, 674-679.

Burger, G., Plante, I., Lonergan, K.M., and Gray, M.W. (1995). The mitochondrial DNA of the amoeboid protozoon, Acanthamoeba castellanii: complete sequence, gene content and genome organization. J Mol Biol *245*, 522-537.

Gott, J.M., Somerlot, B.H., and Gray, M.W. (2010). Two forms of RNA editing are required for tRNA maturation in Physarum mitochondria. RNA *in press*.

Gu, W., Jackman, J.E., Lohan, A.J., Gray, M.W., and Phizicky, E.M. (2003). tRNAHis maturation: an essential yeast protein catalyzes addition of a guanine nucleotide to the 5' end of tRNAHis. Genes & development *17*, 2889-2901.

Holm, P.S., and Krupp, G. (1992). The acceptor stem in pre-tRNAs determines the cleavage specificity of RNase P. Nucleic Acids Res *20*, 421-423.

Hyde, S.J., Eckenroth, B.E., Smith, B.A., Eberley, W.A., Heintz, N.H., Jackman, J.E., and Doublie, S. (2010). tRNA(His) guanylyltransferase (THG1), a unique 3'-5' nucleotidyl transferase, shares unexpected structural homology with canonical 5'-3' DNA polymerases. Proc Natl Acad Sci U S A *107*, 20305-20310.

Jackman, J.E., Gott, J.M., and Gray, M.W. (2012). Doing it in reverse: 3'-to-5' polymerization by the Thg1 superfamily. Rna *18*, 886-899.

Jackman, J.E., and Phizicky, E.M. (2006a). tRNAHis guanylyltransferase adds G-1 to the 5' end of tRNAHis by recognition of the anticodon, one of several features unexpectedly shared with tRNA synthetases. Rna *12*, 1007-1014.

Jackman, J.E., and Phizicky, E.M. (2006b). tRNAHis guanylyltransferase catalyzes a 3'-5' polymerization reaction that is distinct from G-1 addition. Proc Natl Acad Sci U S A *103*, 8640-8645.

Lonergan, K.M., and Gray, M.W. (1993). Editing of transfer RNAs in Acanthamoeba castellanii mitochondria. Science *259*, 812-816.

Marck, C., and Grosjean, H. (2002). tRNomics: Analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. RNA *8*, 1189-1232.

Price, D.H., and Gray, M.W. (1999a). Confirmation of predicted edits and demonstration of unpredicted edits in Acanthamoeba castellanii mitochondrial tRNAs. Current genetics *35*, 23-29.

Price, D.H., and Gray, M.W. (1999b). A novel nucleotide incorporation activity implicated in the editing of mitochondrial transfer RNAs in Acanthamoeba castellanii. Rna *5*, 302-317.

Randau, L., Munch, R., Hohn, M.J., Jahn, D., and Soll, D. (2005). Nanoarchaeum equitans creates functional tRNAs from separate genes for their 5'- and 3'-halves. Nature *433*, 537-541.

Rao, B.S., Maris, E.L., and Jackman, J.E. (2011). tRNA 5'-end repair activities of tRNAHis guanylyltransferase (Thg1)-like proteins from Bacteria and Archaea. Nucleic Acids Research *39*, 1833-1842.

Smith, B.A., and Jackman, J.E. (2012). Kinetic analysis of 3'-5' nucleotide addition catalyzed by eukaryotic tRNA(His) guanylyltransferase. Biochemistry *51*, 453-465.

Wang, C., Sobral, B.W., and Williams, K.P. (2007). Loss of a universal tRNA feature. J Bacteriol *189*, 1954-1962.

Yuan, J., Gogakos, T., Babina, A.M., Soll, D., and Randau, L. (2011). Change of tRNA identity leads to a divergent orthogonal histidyl-tRNA synthetase/tRNAHis pair. Nucleic Acids Res *39*, 2286-2293.